FIG. 1A

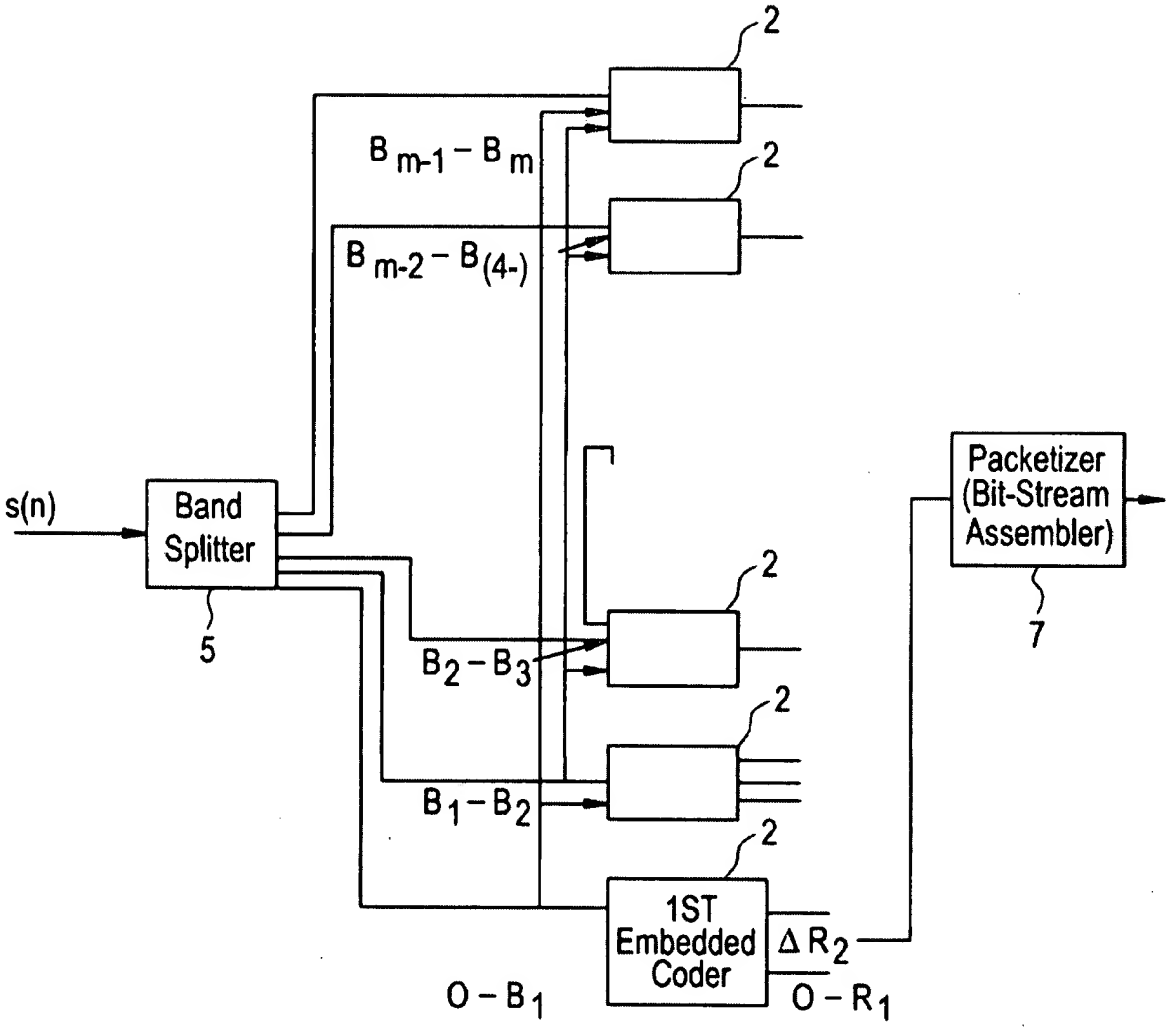

2
2
$B_{m-1} - B_m$
$B_{m-2} - B_{(4-)}$
s(n)
Band Splitter
5
2
$B_2 - B_3$
2
$B_1 - B_2$
2
1ST Embedded Coder
ΔR_2
$O - B_1$
$O - R_1$
Packetizer (Bit-Stream Assembler)
7

FIG. 1B

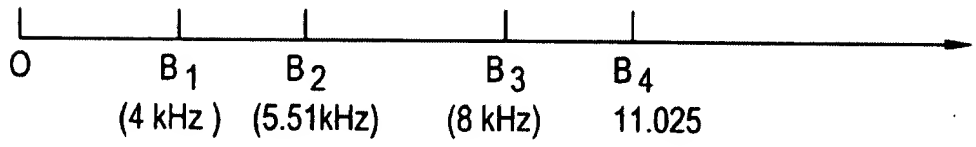

O
B_1
(4 kHz)
B_2
(5.51kHz)
B_3
(8 kHz)
B_4
11.025

FIG. 2A

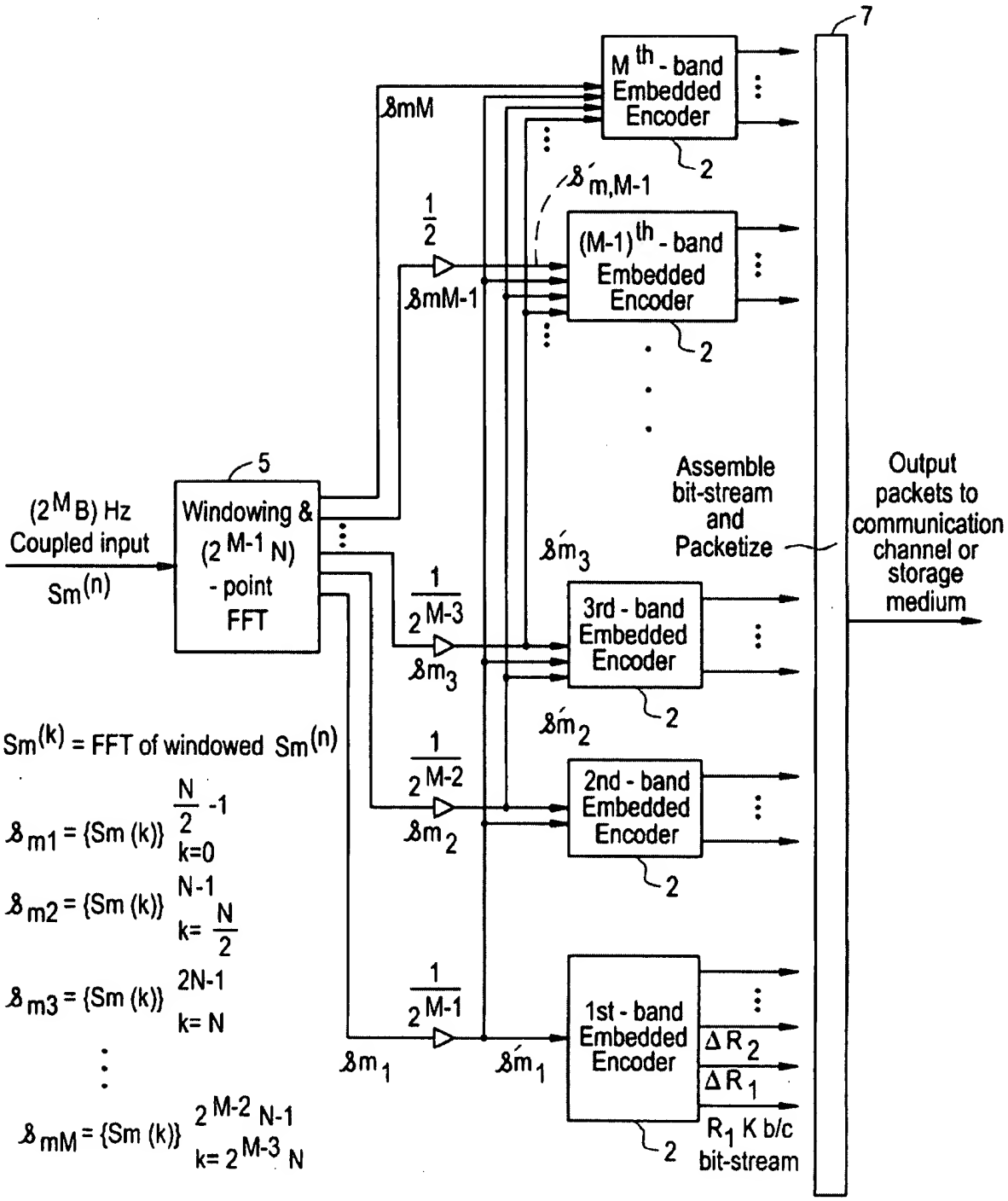

7
M th - band Embedded Encoder
2
$\mathcal{S}$mM
$\mathcal{S}'_{m,M-1}$
1/2
(M-1) th - band Embedded Encoder
2
$\mathcal{S}$mM-1
5
(2 M B) Hz Coupled input
Sm(n)
Windowing & (2 M-1 N) - point FFT
Assemble bit-stream and Packetize
Output packets to communication channel or storage medium
$\mathcal{S}'_{m3}$
1/2 M-3
3rd - band Embedded Encoder
2
$\mathcal{S}$m3
$\mathcal{S}'_{m2}$
1/2 M-2
2nd - band Embedded Encoder
2
$\mathcal{S}$m2
Sm(k) = FFT of windowed Sm(n)
$\mathcal{S}_{m1}$ = {Sm (k)} N/2 -1, k=0
$\mathcal{S}_{m2}$ = {Sm (k)} N-1, k= N/2
$\mathcal{S}_{m3}$ = {Sm (k)} 2N-1, k= N
$\mathcal{S}_{mM}$ = {Sm (k)} 2 M-2 N-1, k= 2 M-3 N
1/2 M-1
1st - band Embedded Encoder
2
$\mathcal{S}$m1
$\mathcal{S}'_{m1}$
ΔR2
ΔR1
R1 K b/c bit-stream

FIG. 2B

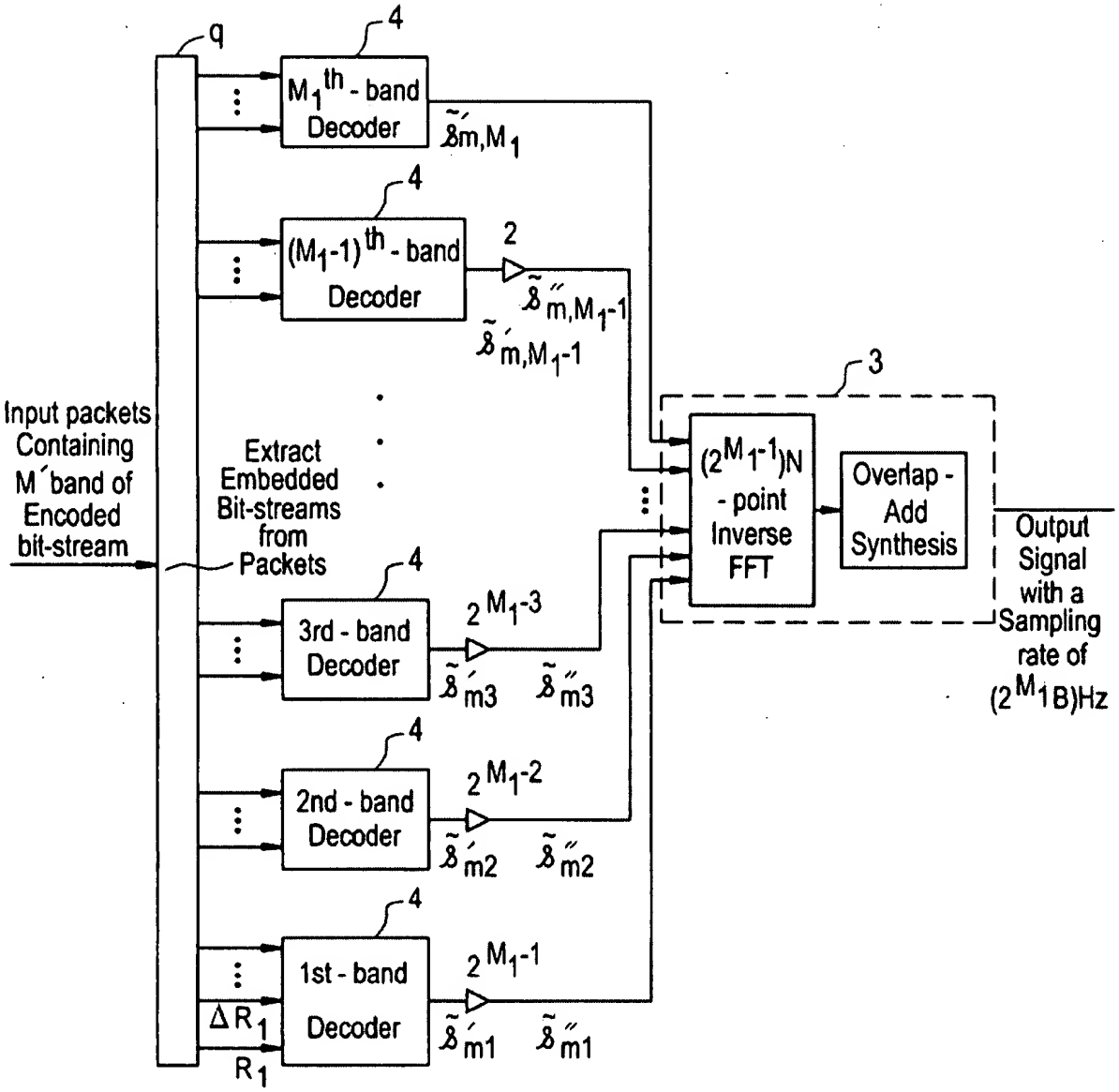

q
4
M_1^{th} - band Decoder
$\tilde{\hat{s}}_{m,M_1}$
4
$(M_1-1)^{th}$ - band Decoder
2
$\tilde{s}''_{m,M_1-1}$
$\tilde{s}'_{m,M_1-1}$
3
$(2^{M_1-1})N$ - point Inverse FFT
Overlap - Add Synthesis
Output Signal with a Sampling rate of $(2^{M_1}B)$Hz
Input packets Containing M´band of Encoded bit-stream
Extract Embedded Bit-streams from Packets
4
3rd - band Decoder
2^{M_1-3}
$\tilde{s}'_{m3}$
$\tilde{s}''_{m3}$
4
2nd - band Decoder
2^{M_1-2}
$\tilde{s}'_{m2}$
$\tilde{s}''_{m2}$
4
1st - band Decoder
2^{M_1-1}
$\tilde{s}'_{m1}$
$\tilde{s}''_{m1}$
ΔR_1
R_1

FIG. 3A

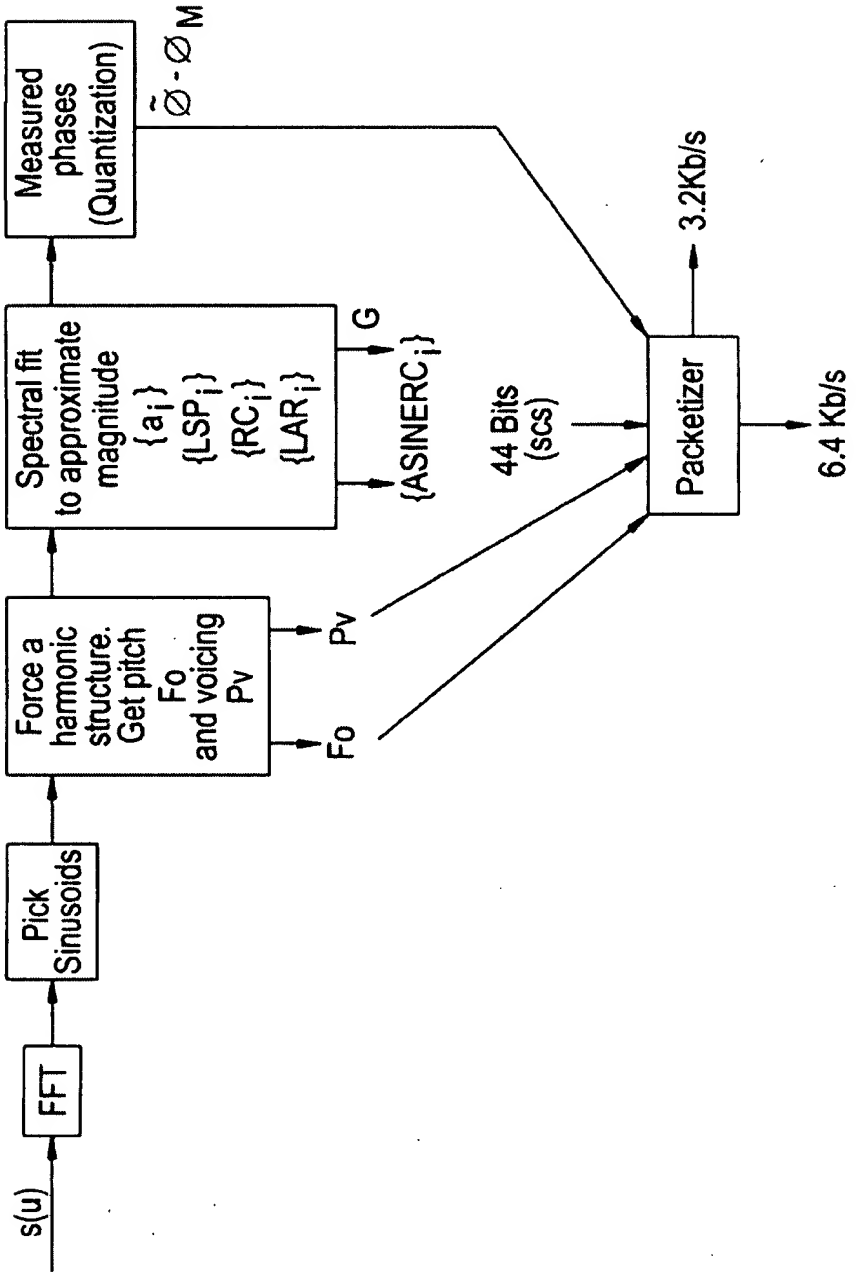

s(u)
FFT
Pick Sinusoids
Force a harmonic structure. Get pitch Fo and voicing Pv
Fo
Pv
Spectral fit to approximate magnitude {a_i} {LSP_i} {RC_i} {LAR_i}
G
{ASINERC_i}
Measured phases (Quantization)
Ø̃ - Ø_M
44 Bits (scs)
Packetizer
3.2Kb/s
6.4 Kb/s

FIG. 3B

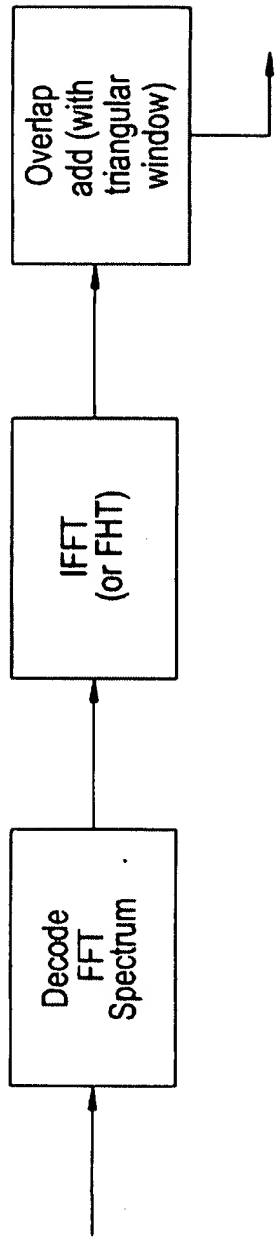

Overlap add (with triangular window)
IFFT (or FHT)
Decode FFT Spectrum
Decoder: Synthesis every M ms.

FIG. 4A

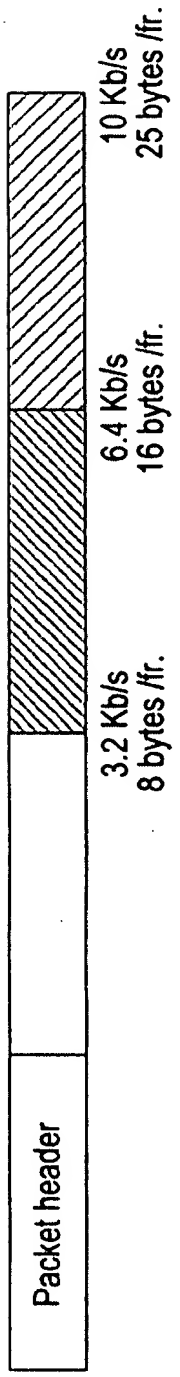

Bit Stream
Packet header
3.2 Kb/s
8 bytes /fr.
6.4 Kb/s
16 bytes /fr.
10 Kb/s
25 bytes /fr.

FIG. 4B

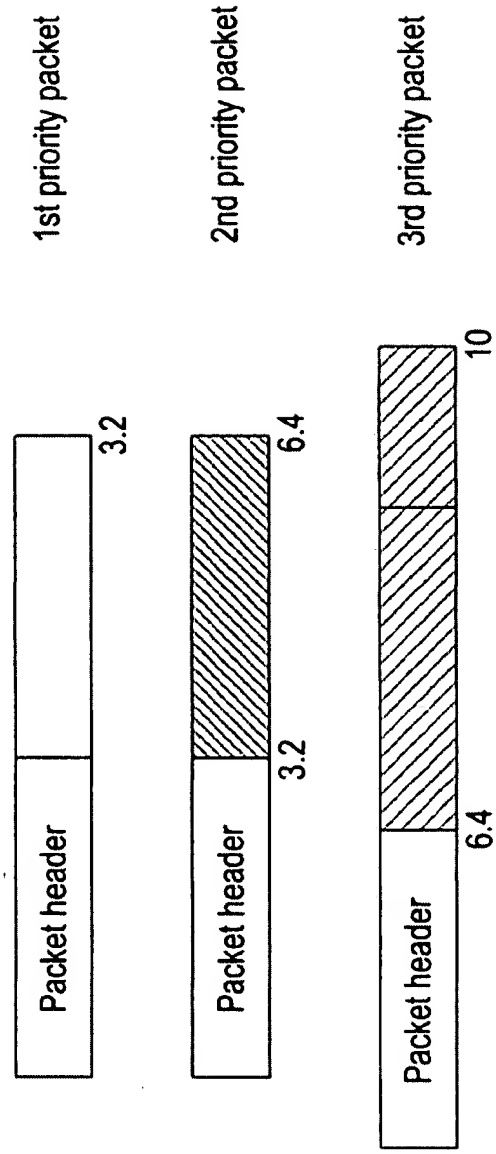

1st priority packet
2nd priority packet
3rd priority packet
Packet header
Packet header
Packet header
3.2
6.4
10
3.2
6.4

FIG. 5

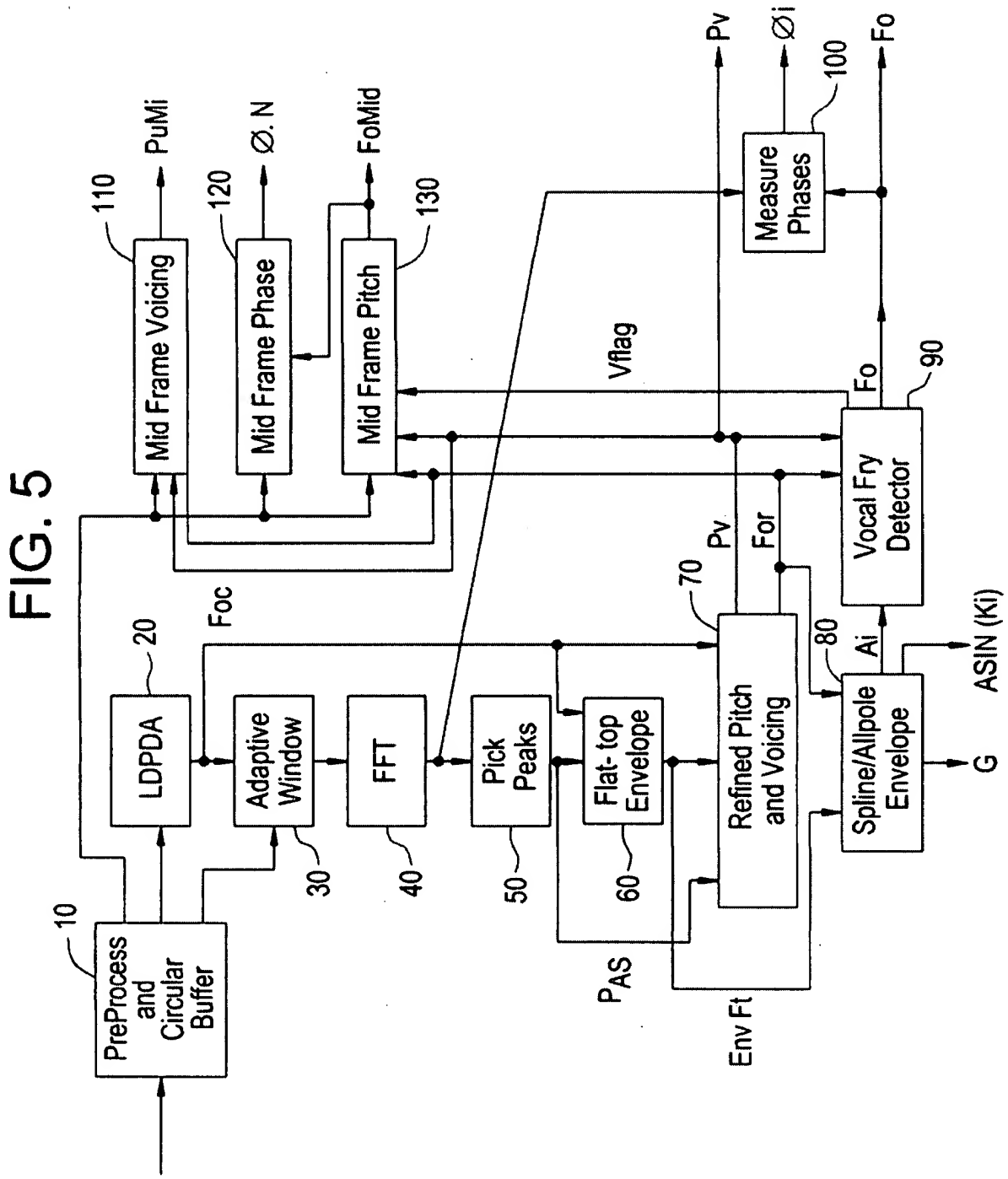

PreProcess and Circular Buffer
10
LDPDA
20
Adaptive Window
30
FFT
40
Pick Peaks
50
Flat- top Envelope
60
Refined Pitch and Voicing
70
Spline/Allpole Envelope
80
Vocal Fry Detector
90
Measure Phases
100
Mid Frame Voicing
110
Mid Frame Phase
120
Mid Frame Pitch
130
PuMi
Ø. N
FoMid
Foc
Vflag
Pv
Øi
Fo
For
Ai
G
ASIN (Ki)
PAS
Env Ft

FIG. 5A

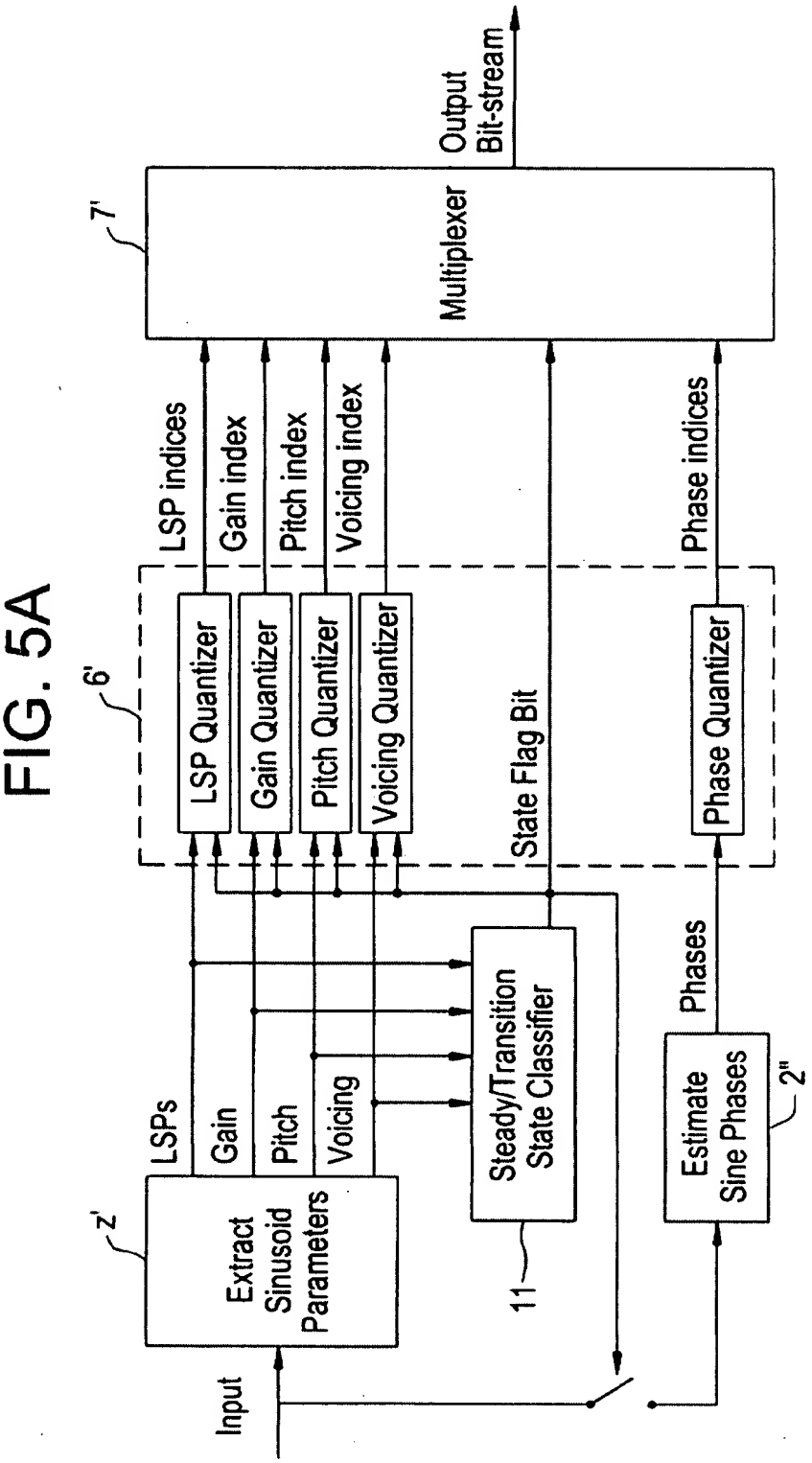

Output Bit-stream
7'
Multiplexer
LSP indices
Gain index
Pitch index
Voicing index
Phase indices
6'
LSP Quantizer
Gain Quantizer
Pitch Quantizer
Voicing Quantizer
State Flag Bit
Phase Quantizer
LSPs
Gain
Pitch
Voicing
Phases
5'
Extract Sinusoid Parameters
Steady/Transition State Classifier
11
Estimate Sine Phases
2"
Input

FIG. 6

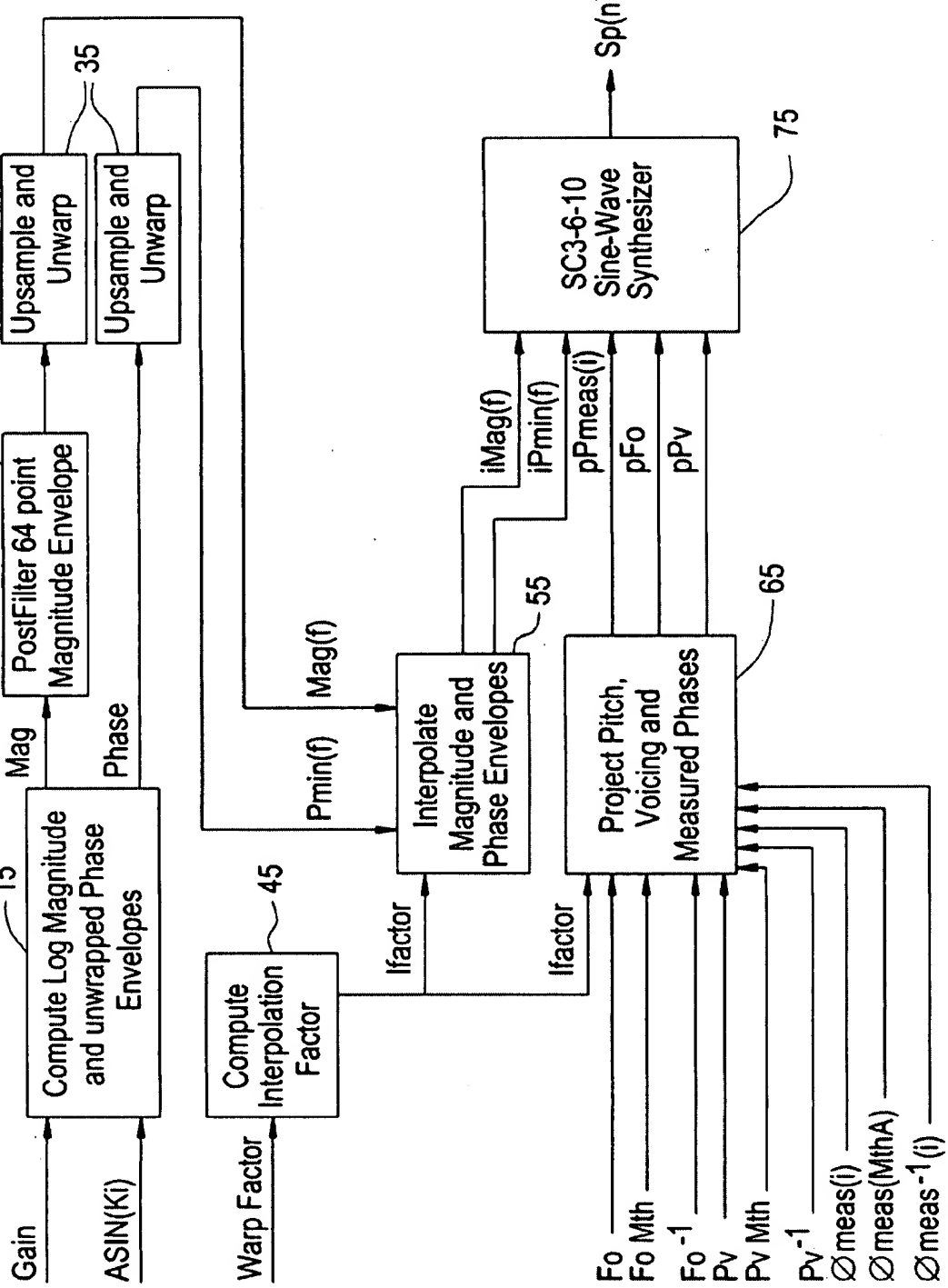

Gain
ASIN(Ki)
Warp Factor
Compute Log Magnitude and unwrapped Phase Envelopes
15
Compute Interpolation Factor
45
Mag
Phase
PostFilter 64 point Magnitude Envelope
25
Upsample and Unwarp
Upsample and Unwarp
35
Pmin(f)
Mag(f)
Interpolate Magnitude and Phase Envelopes
55
Ifactor
Ifactor
iMag(f)
iPmin(f)
pPmeas(i)
pFo
pPv
Project Pitch, Voicing and Measured Phases
65
SC3-6-10 Sine-Wave Synthesizer
75
Sp(n)
Fo
Fo Mth
Fo-1
Pv
Pv Mth
Pv-1
Ømeas(i)
Ømeas(MthA)
Ømeas-1(i)

FIG. 6A

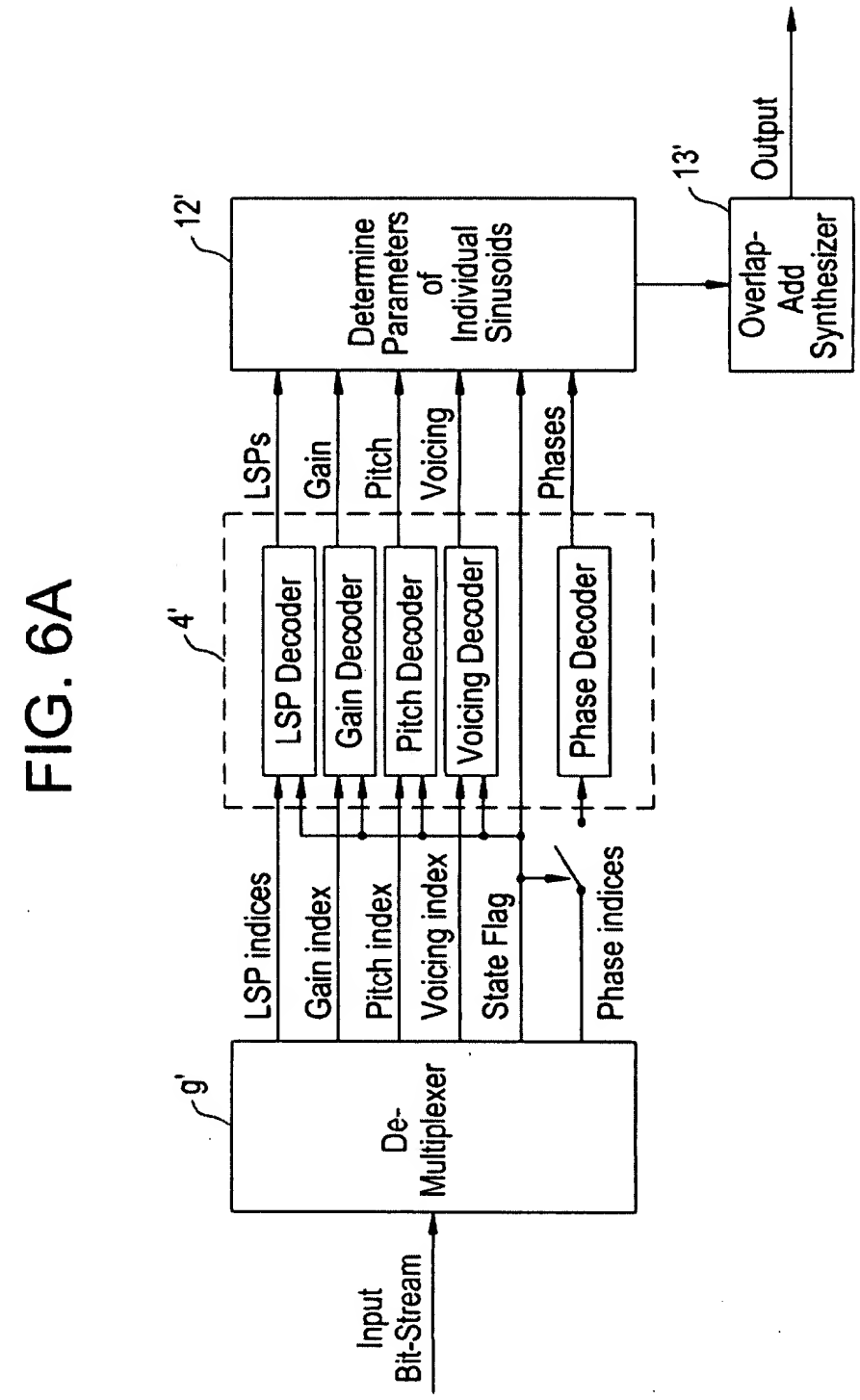

Input Bit-Stream
9'
De-Multiplexer
LSP indices
Gain index
Pitch index
Voicing index
State Flag
Phase indices
4'
LSP Decoder
Gain Decoder
Pitch Decoder
Voicing Decoder
Phase Decoder
LSPs
Gain
Pitch
Voicing
Phases
12'
Determine Parameters of Individual Sinusoids
13'
Overlap-Add Synthesizer
Output

FIG. 7

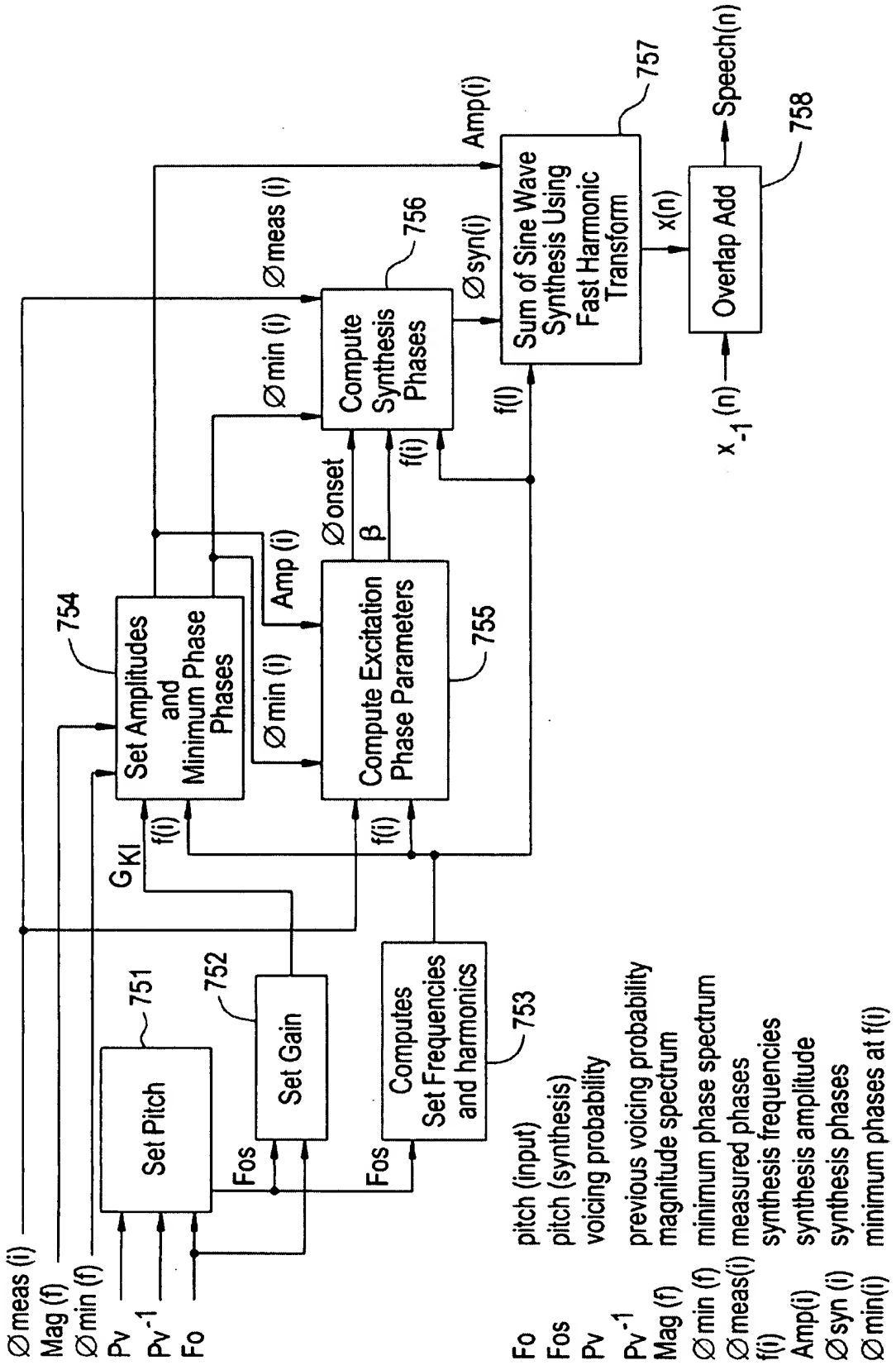

Ø meas (i)
Mag (f)
Ø min (f)
Pv
Pv-1
Fo
Set Pitch
751
Fos
Set Gain
752
Fos
Computes Set Frequencies and harmonics
753
GKI
f(i)
Set Amplitudes and Minimum Phase Phases
754
Ø min (i)
Amp (i)
Ø min (i)
Ø onset
β
f(i)
Compute Excitation Phase Parameters
755
Ø meas (i)
Ø min (i)
f(i)
Compute Synthesis Phases
756
Amp(i)
Ø syn(i)
f(I)
Sum of Sine Wave Synthesis Using Fast Harmonic Transform
757
x(n)
Overlap Add
758
Speech(n)
x-1 (n)
Fo pitch (input)
Fos pitch (synthesis)
Pv voicing probability
Pv-1 previous voicing probability
Mag (f) magnitude spectrum
Ø min (f) minimum phase spectrum
Ø meas(i) measured phases
f(i) synthesis frequencies
Amp(i) synthesis amplitude
Ø syn (i) synthesis phases
Ø min(i) minimum phases at f(i)

FIG. 8

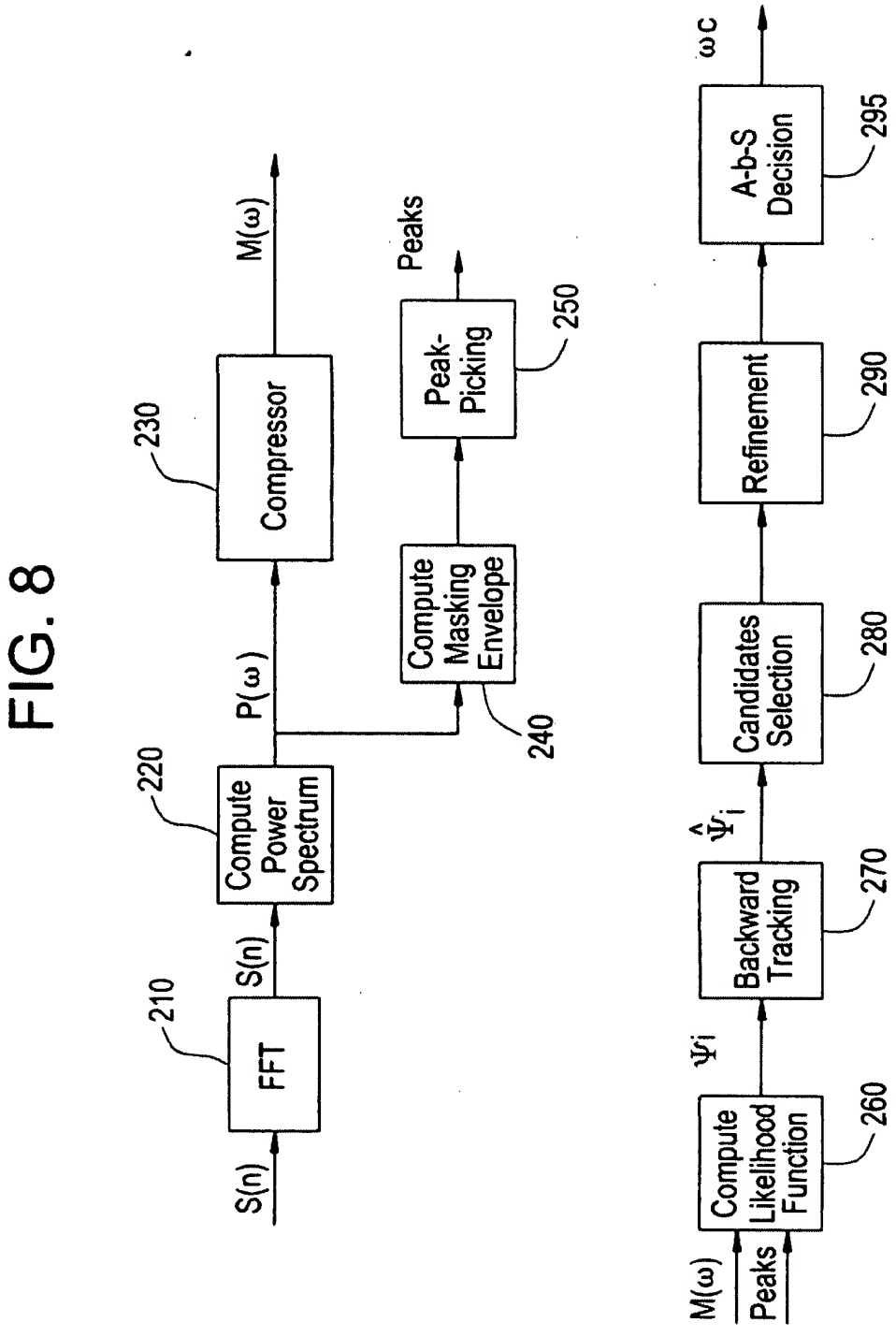

S(n)
FFT
210
S(n)
Compute Power Spectrum
220
P(ω)
Compressor
230
M(ω)
Compute Masking Envelope
240
Peak-Picking
250
Peaks
Compute Likelihood Function
260
Ψi
Backward Tracking
270
Ψ̂i
Candidates Selection
280
Refinement
290
A-b-S Decision
295
ωc
M(ω)
Peaks

FIG. 8A

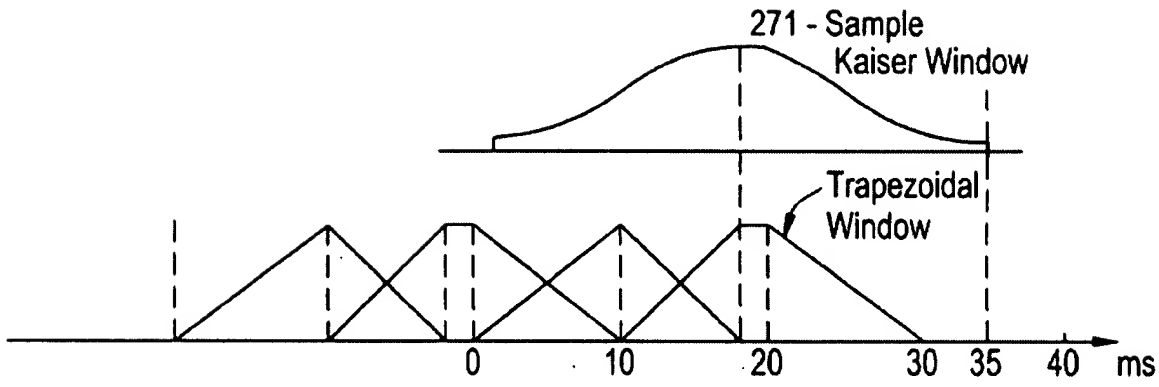

271 - Sample
Kaiser Window
Trapezoidal
Window
0
10
20
30
35
40
ms

FIG. 9A

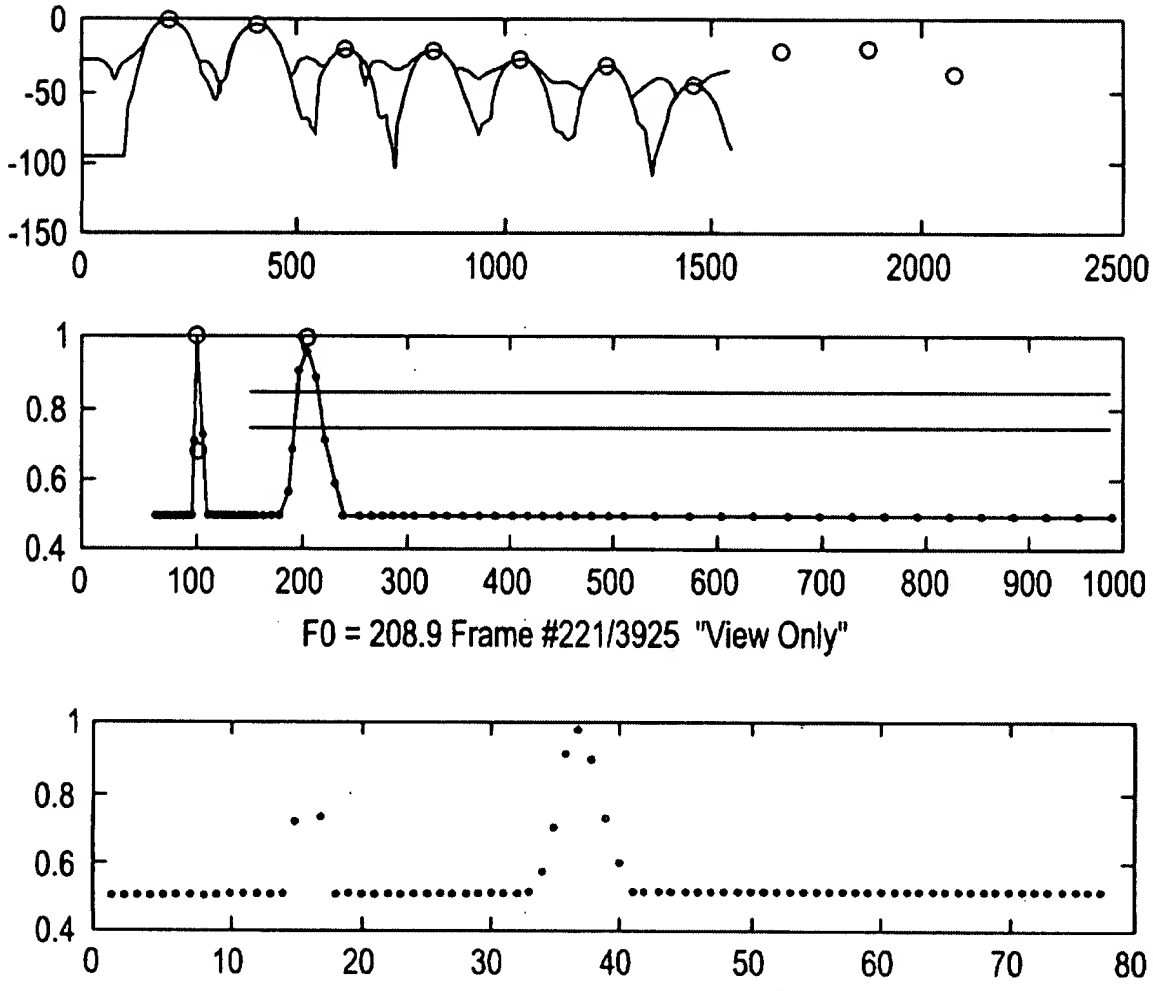

0
-50
-100
-150
0
500
1000
1500
2000
2500
1
0.8
0.6
0.4
0
100
200
300
400
500
600
700
800
900
1000
F0 = 208.9 Frame #221/3925 "View Only"
1
0.8
0.6
0.4
0
10
20
30
40
50
60
70
80

FIG. 9B

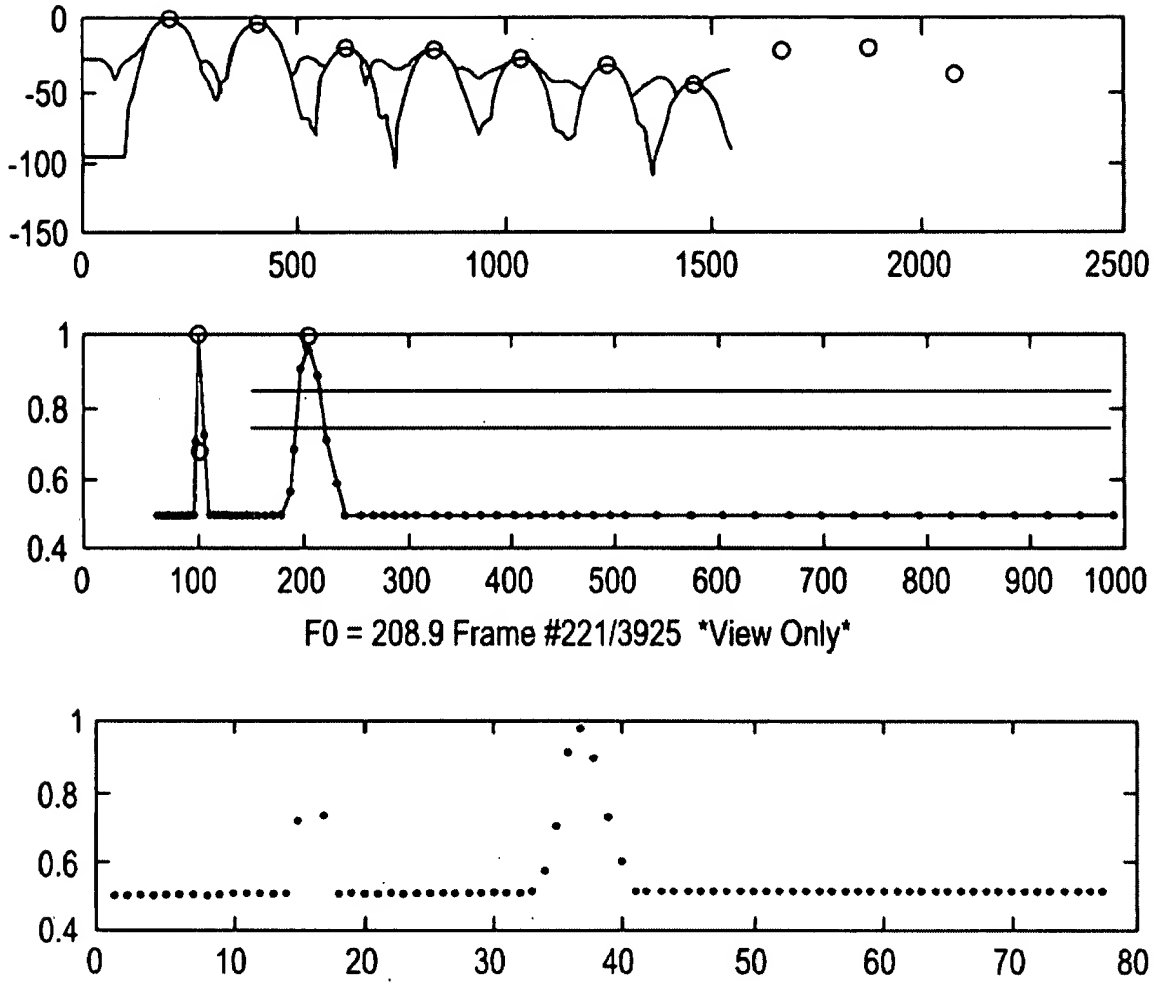

0
-50
-100
-150
0
500
1000
1500
2000
2500
1
0.8
0.6
0.4
0
100
200
300
400
500
600
700
800
900
1000
F0 = 208.9 Frame #221/3925 *View Only*
1
0.8
0.6
0.4
0
10
20
30
40
50
60
70
80

FIG. 9C

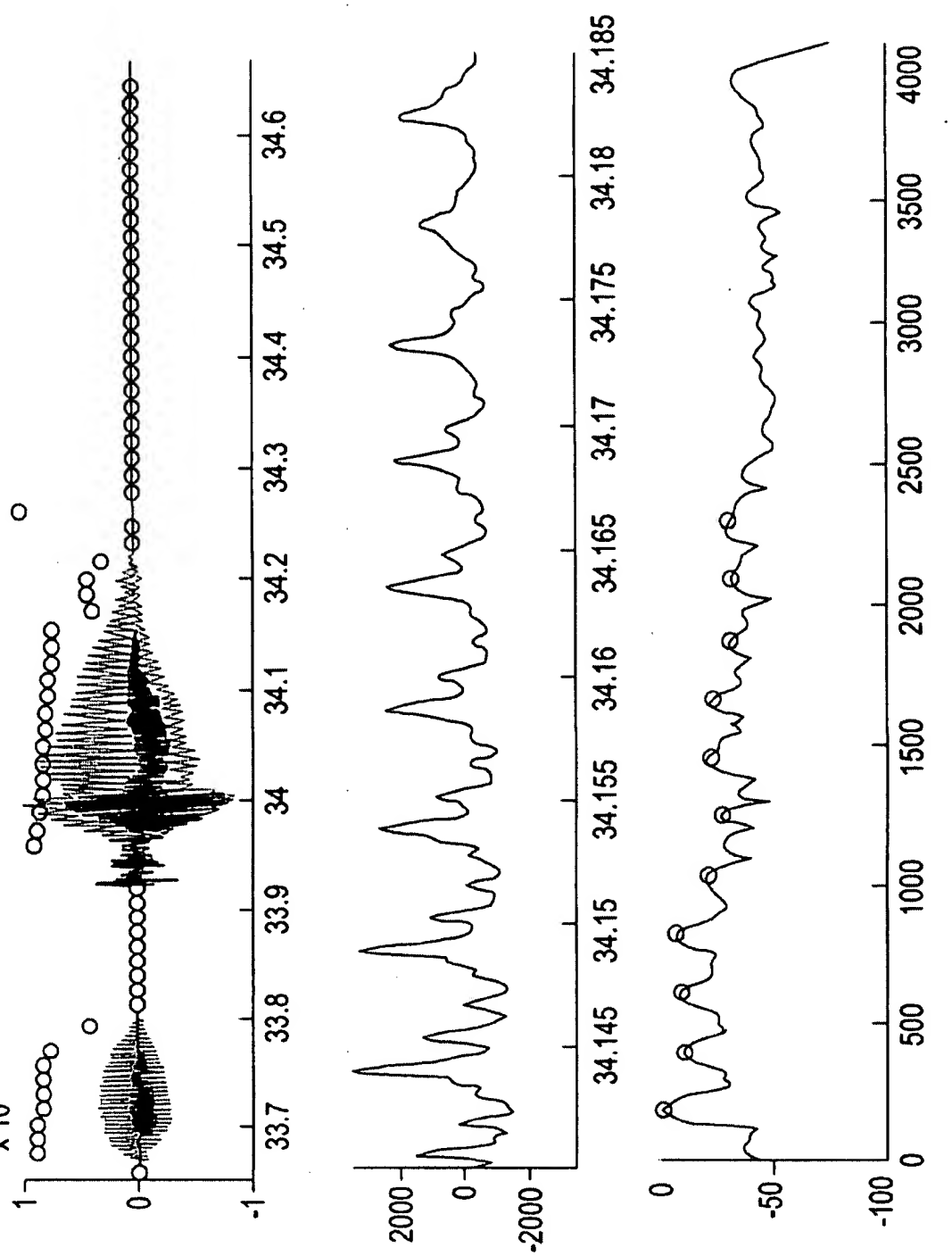

F0 = 203.9 Frame #2278/3925 "Views Only"
x 10^4
1
0
-1
33.7
33.8
33.9
34
34.1
34.2
34.3
34.4
34.5
34.6
2000
0
-2000
34.145
34.15
34.155
34.16
34.165
34.17
34.175
34.18
34.185
0
-50
-100
0
500
1000
1500
2000
2500
3000
3500
4000

FIG. 9D

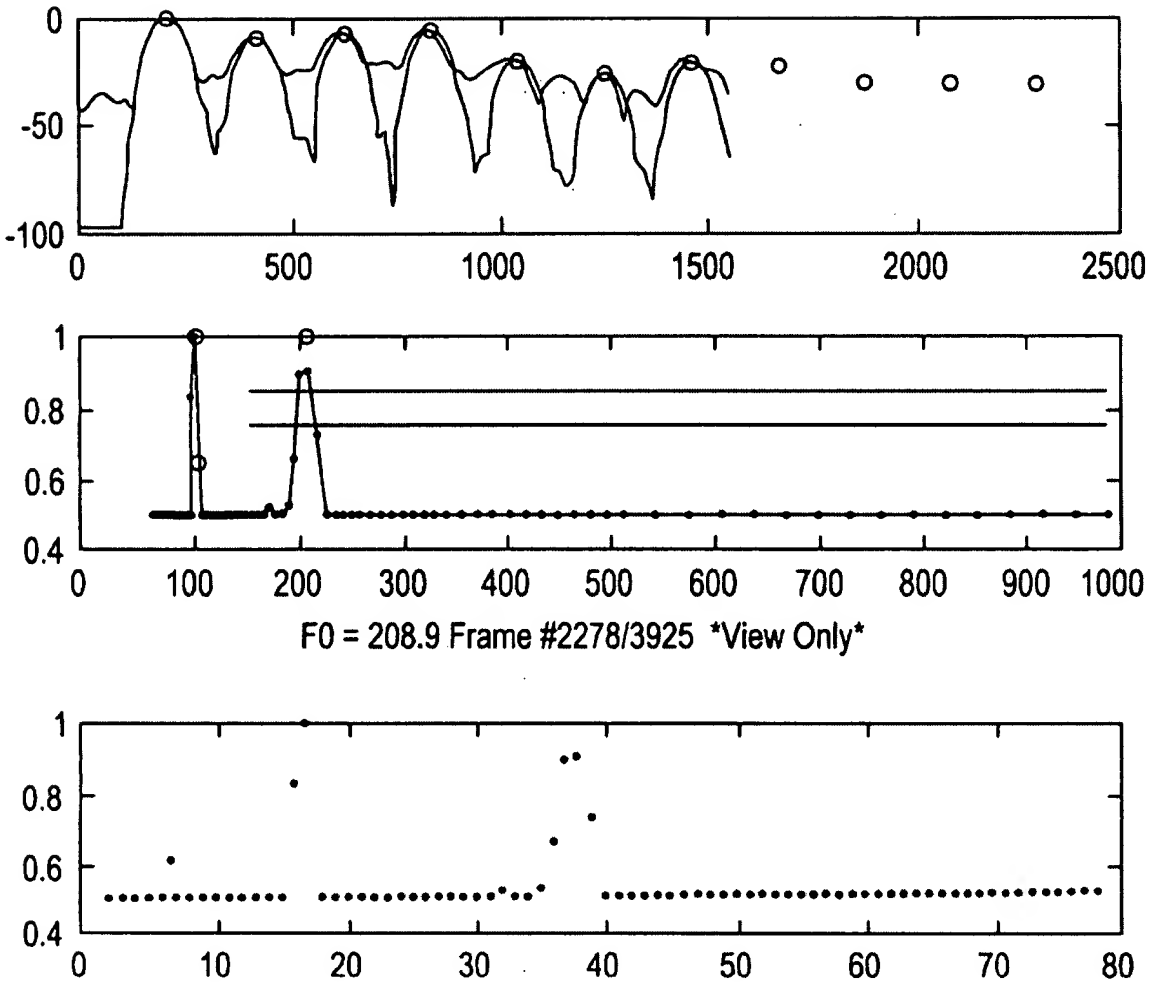

0
-50
-100
0
500
1000
1500
2000
2500
1
0.8
0.6
0.4
0
100
200
300
400
500
600
700
800
900
1000
F0 = 208.9 Frame #2278/3925 *View Only*
1
0.8
0.6
0.4
0
10
20
30
40
50
60
70
80

FIG. 10

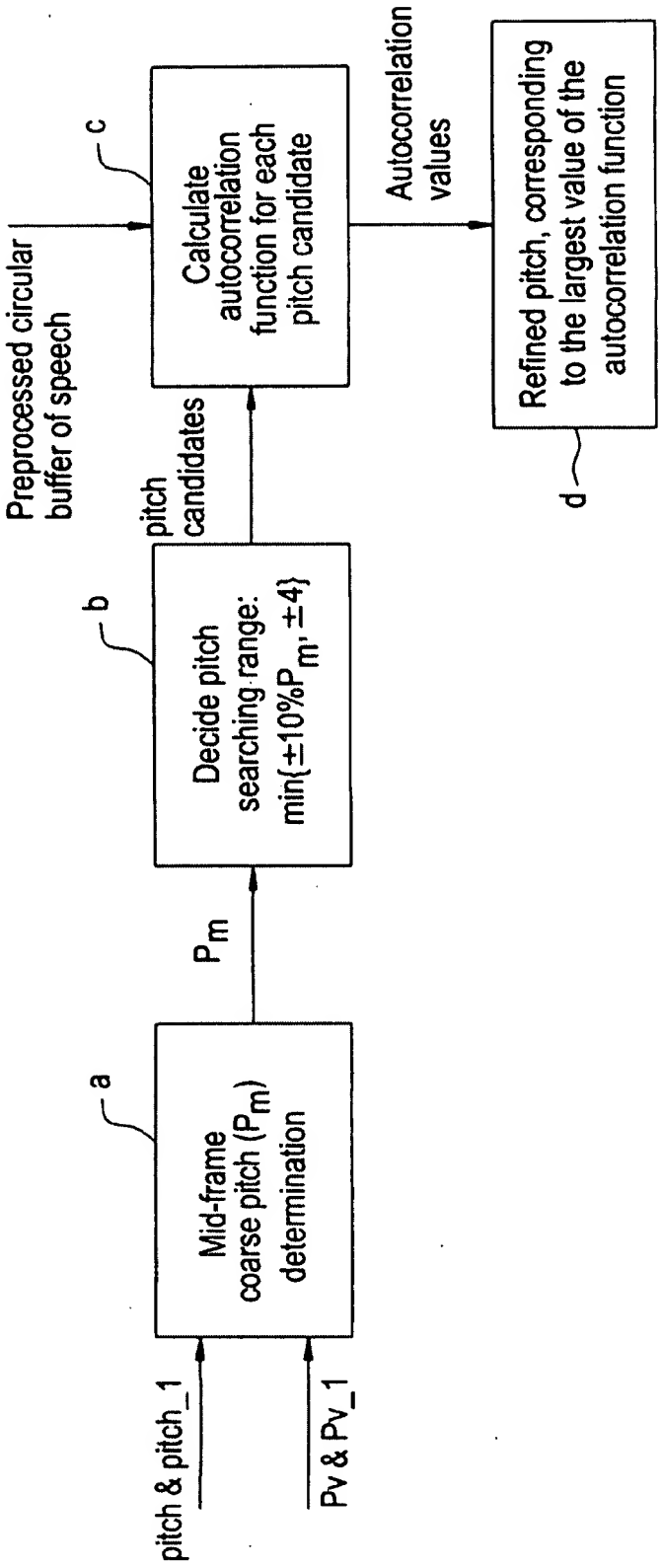

pitch & pitch_1
Pv & Pv_1
a
Mid-frame coarse pitch (P_m) determination
P_m
b
Decide pitch searching range: min{±10%P_m, ±4}
pitch candidates
Preprocessed circular buffer of speech
c
Calculate autocorrelation function for each pitch candidate
Autocorrelation values
Refined pitch, corresponding to the largest value of the autocorrelation function
d

FIG. 11

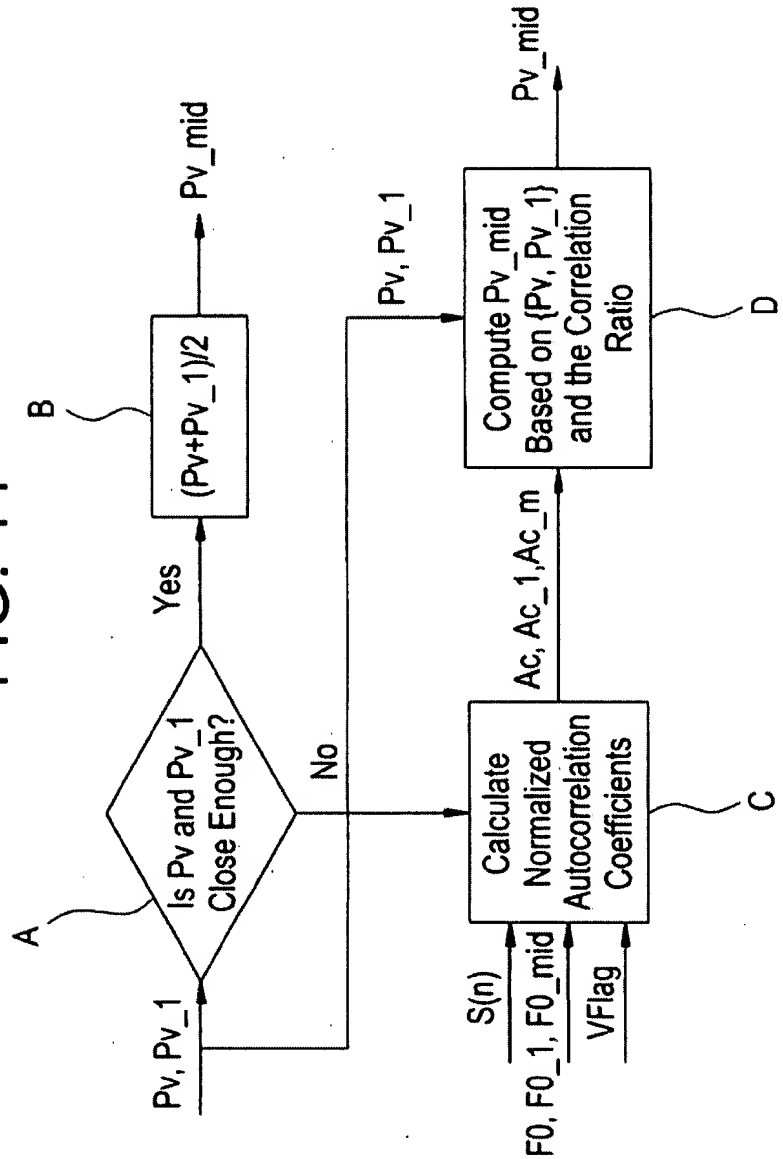

A
Is Pv and Pv_1 Close Enough?
Yes
No
B
(Pv+Pv_1)/2
Pv_mid
Pv, Pv_1
Pv, Pv_1
C
Calculate Normalized Autocorrelation Coefficients
S(n)
F0, F0_1, F0_mid
VFlag
Ac, Ac_1,Ac_m
D
Compute Pv_mid Based on {Pv, Pv_1} and the Correlation Ratio
Pv_mid

FIG. 12

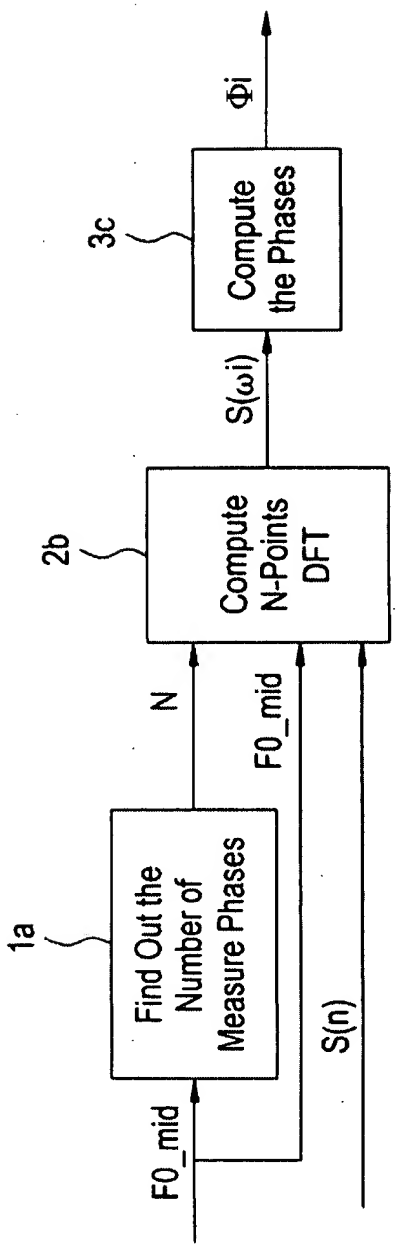

1a
Find Out the Number of Measure Phases
2b
Compute N-Points DFT
3c
Compute the Phases
F0_mid
N
F0_mid
S(n)
S(ωi)
Φi

FIG. 13

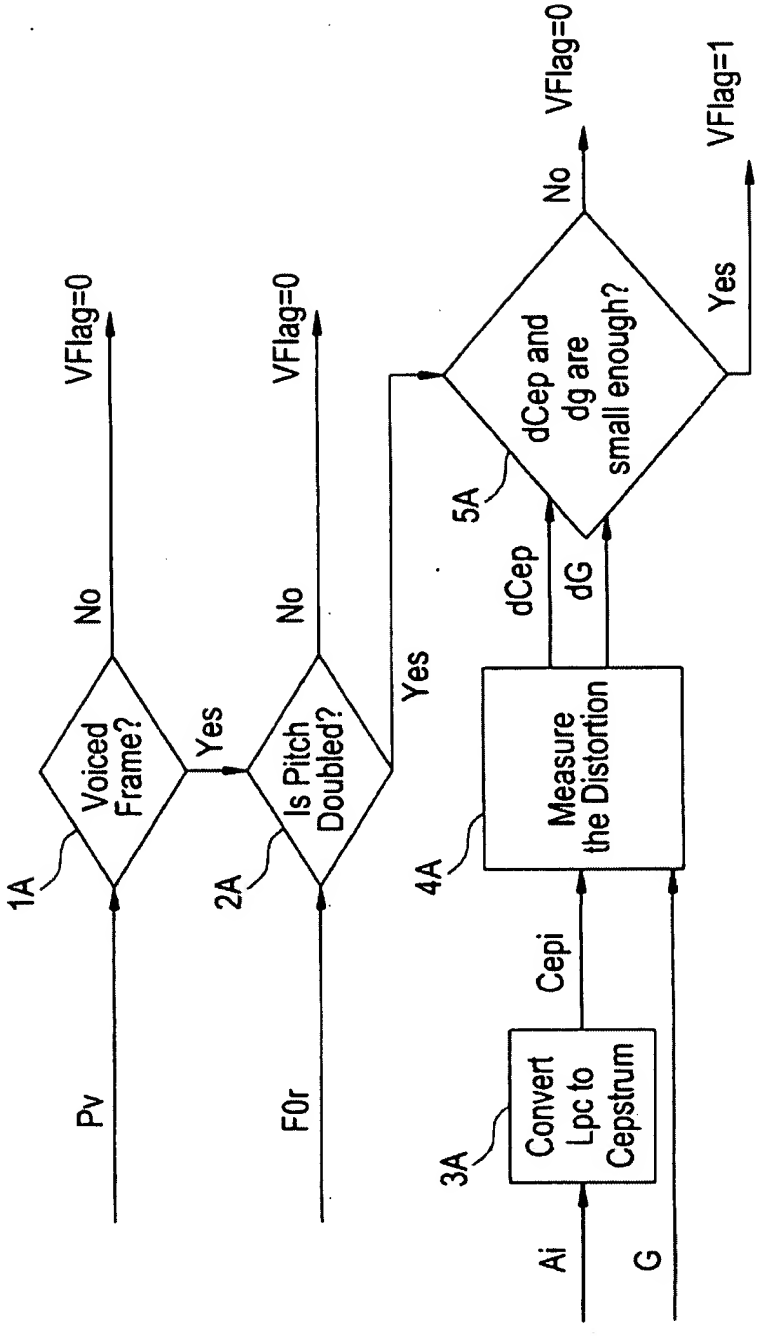

Voiced Frame?
1A
No
VFlag=0
Pv
Yes
Is Pitch Doubled?
2A
No
VFlag=0
F0r
Yes
Convert Lpc to Cepstrum
3A
Ai
Cepi
G
Measure the Distortion
4A
dCep
dG
dCep and dg are small enough?
5A
No
VFlag=0
Yes
VFlag=1

FIG. 14

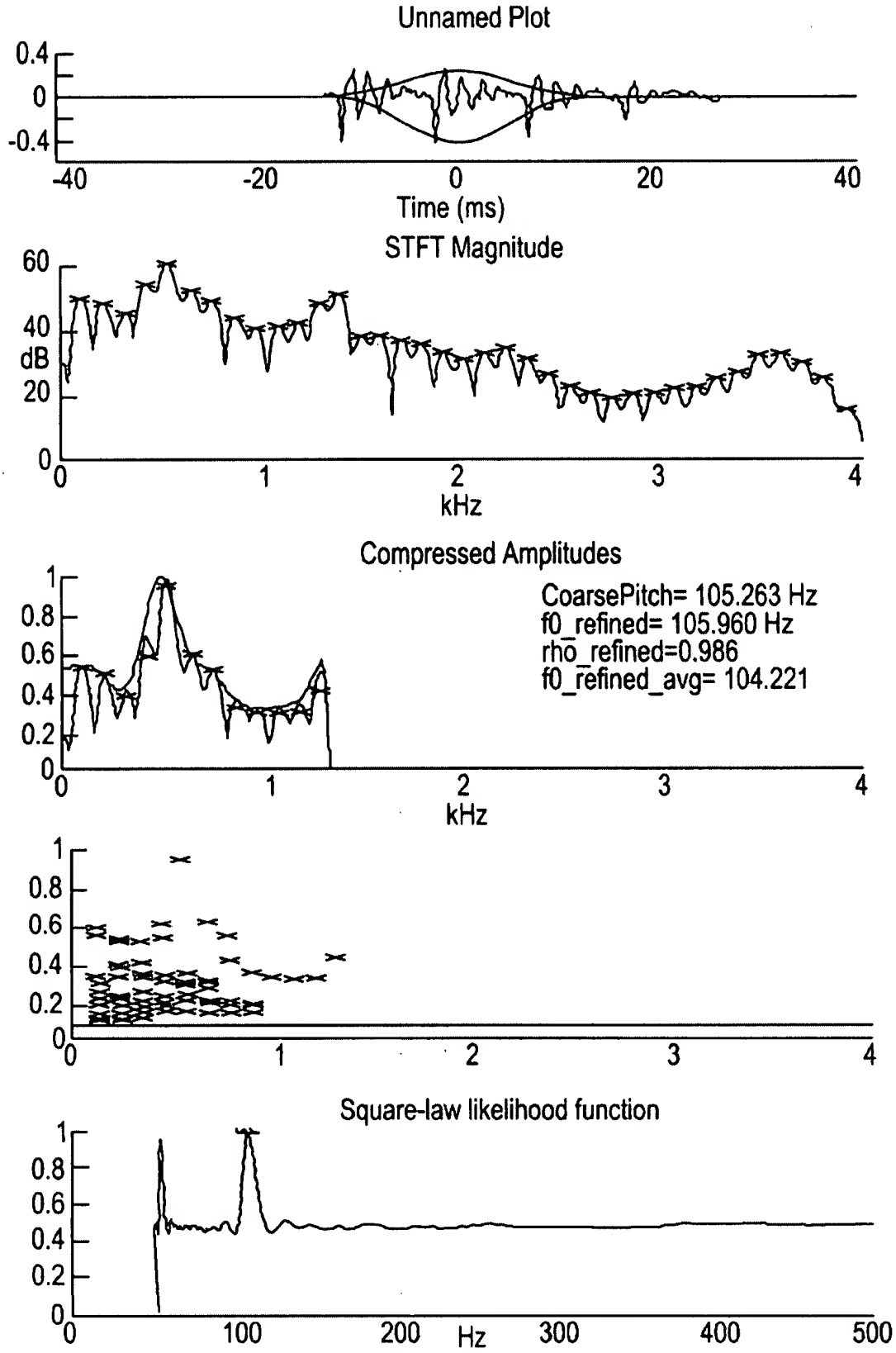

Unnamed Plot
0.4
0
-0.4
-40
-20
0
20
40
Time (ms)
STFT Magnitude
60
40
dB
20
0
0
1
2
3
4
kHz
Compressed Amplitudes
CoarsePitch= 105.263 Hz
f0_refined= 105.960 Hz
rho_refined=0.986
f0_refined_avg= 104.221
1
0.8
0.6
0.4
0.2
0
0
1
2
3
4
kHz
1
0.8
0.6
0.4
0.2
0
0
1
2
3
4
Square-law likelihood function
1
0.8
0.6
0.4
0.2
0
0
100
200
Hz
300
400
500

FIG. 15

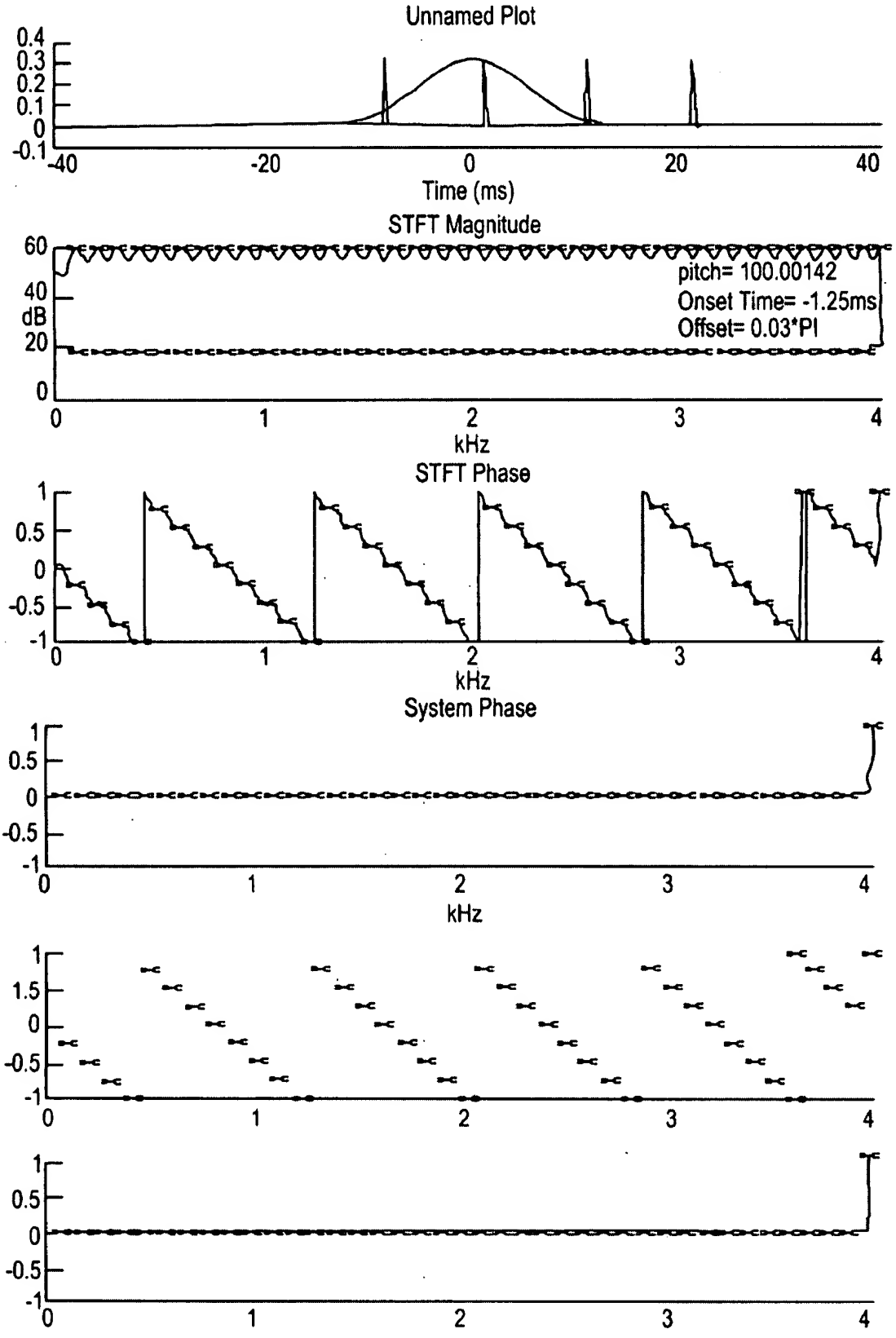

Unnamed Plot
0.4
0.3
0.2
0.1
0
-0.1
-40
-20
0
20
40
Time (ms)
STFT Magnitude
60
40
dB
20
0
pitch= 100.00142
Onset Time= -1.25ms
Offset= 0.03*PI
0
1
2
3
4
kHz
STFT Phase
1
0.5
0
-0.5
-1
0
1
2
3
4
kHz
System Phase
1
0.5
0
-0.5
-1
0
1
2
3
4
kHz
1
1.5
0
-0.5
-1
0
1
2
3
4
1
0.5
0
-0.5
-1
0
1
2
3
4

FIG. 16

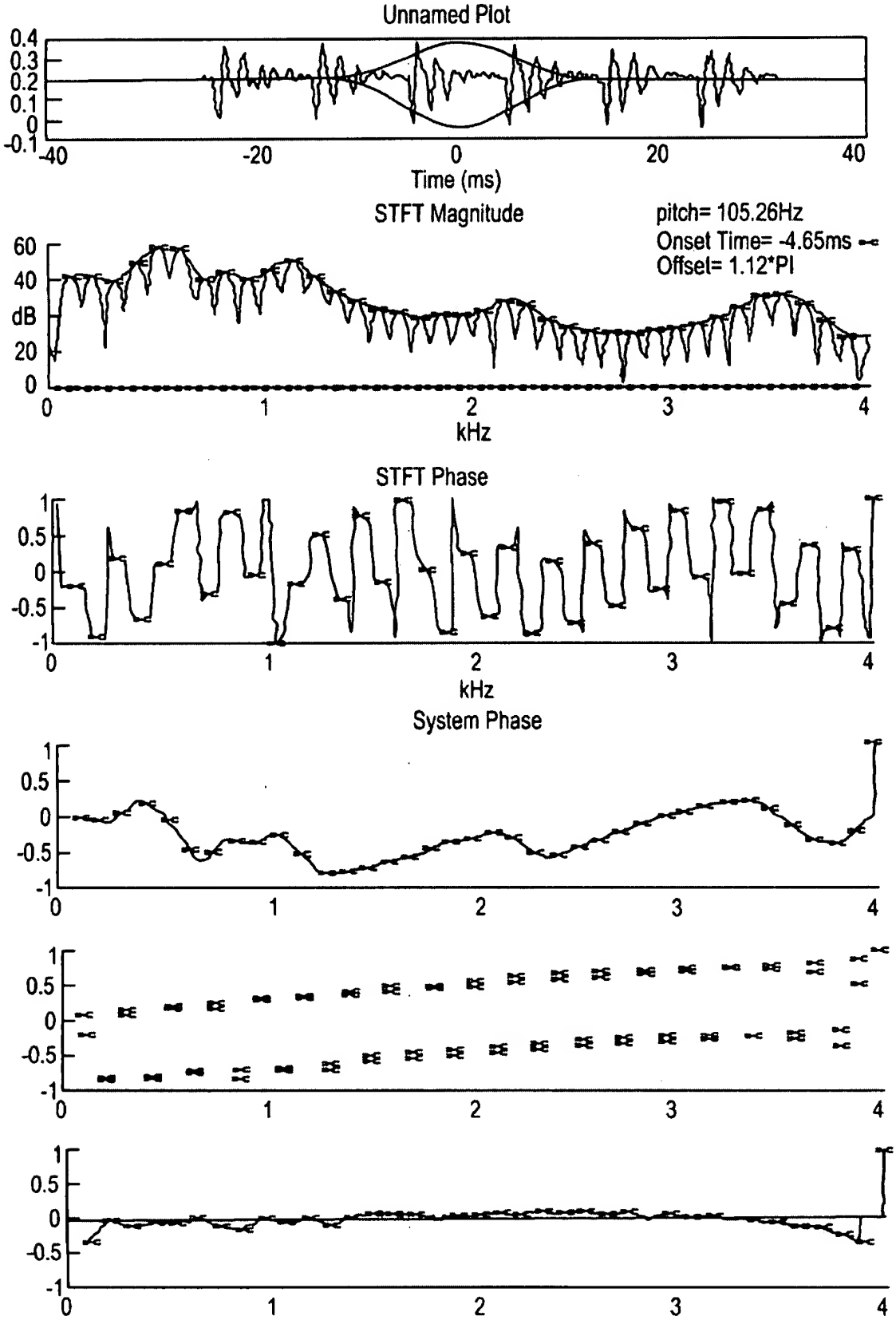

Unnamed Plot
0.4
0.3
0.2
0.1
0
-0.1
-40
-20
0
20
40
Time (ms)
STFT Magnitude
pitch= 105.26Hz
Onset Time= -4.65ms
Offset= 1.12*PI
60
40
dB
20
0
0
1
2
3
4
kHz
STFT Phase
1
0.5
0
-0.5
-1
0
1
2
3
4
kHz
System Phase
1
0.5
0
-0.5
-1
0
1
2
3
4
1
0.5
0
-0.5
-1
0
1
2
3
4
1
0.5
0
-0.5
-1
0
1
2
3
4

FIG. 17

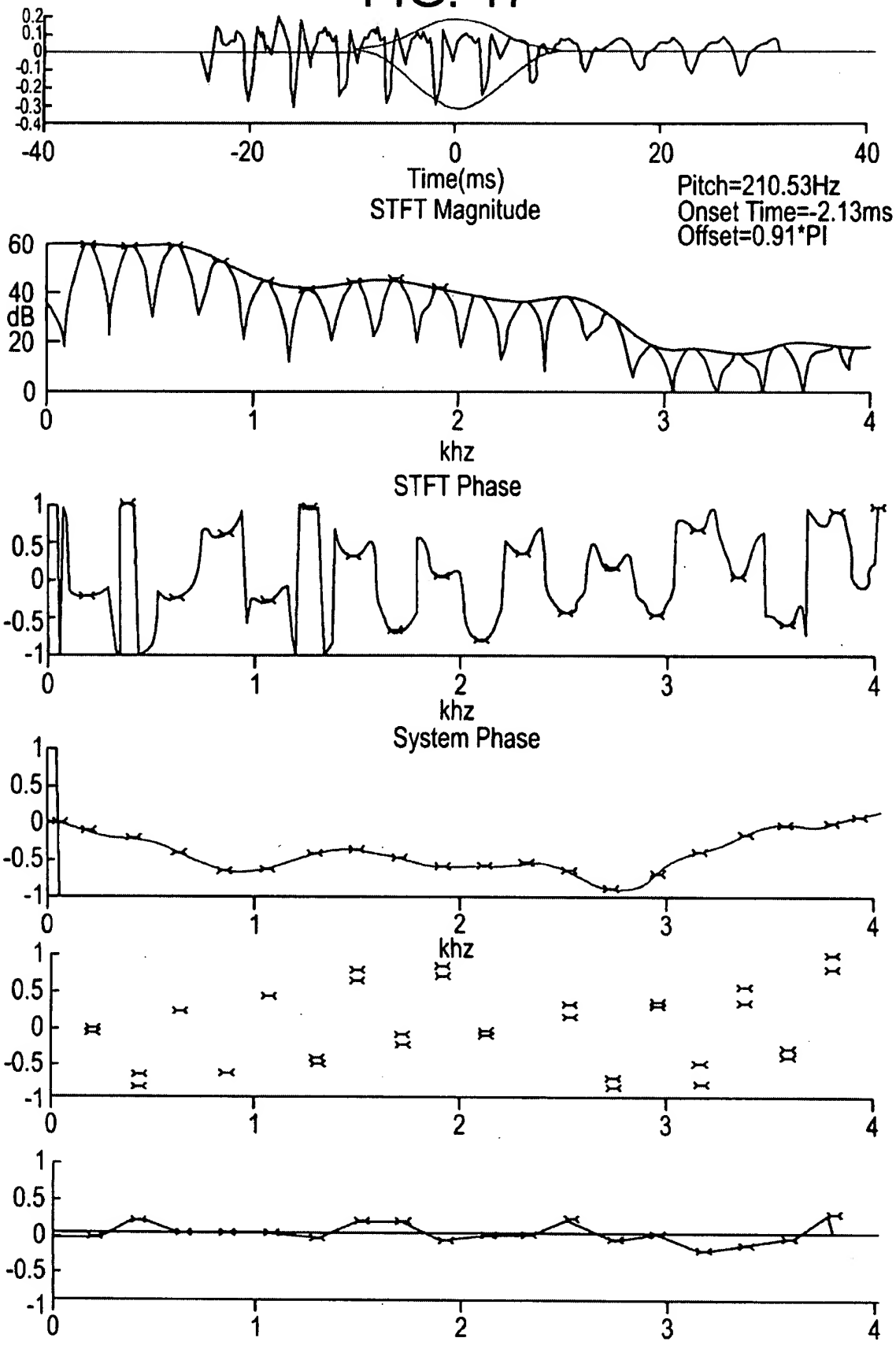

Time(ms)
Pitch=210.53Hz
Onset Time=-2.13ms
Offset=0.91*PI
STFT Magnitude
dB
khz
STFT Phase
khz
System Phase
khz

FIG. 18

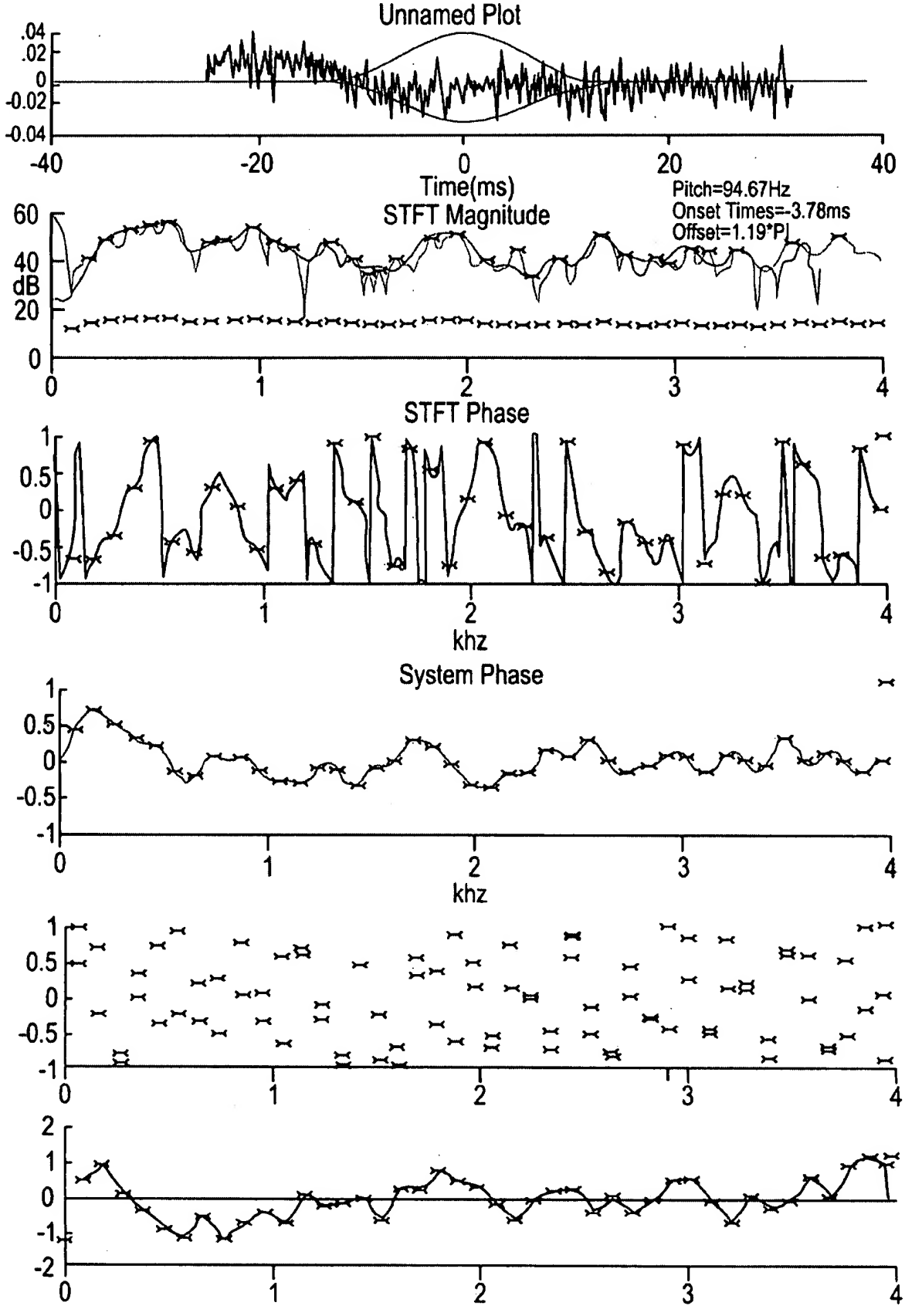

Unnamed Plot
.04
.02
0
-0.02
-0.04
-40
-20
0
20
40
Time(ms)
STFT Magnitude
Pitch=94.67Hz
Onset Times=-3.78ms
Offset=1.19*Pi
60
40
dB
20
0
0
1
2
3
4
STFT Phase
1
0.5
0
-0.5
-1
khz
System Phase
khz
2
-2

FIG. 19

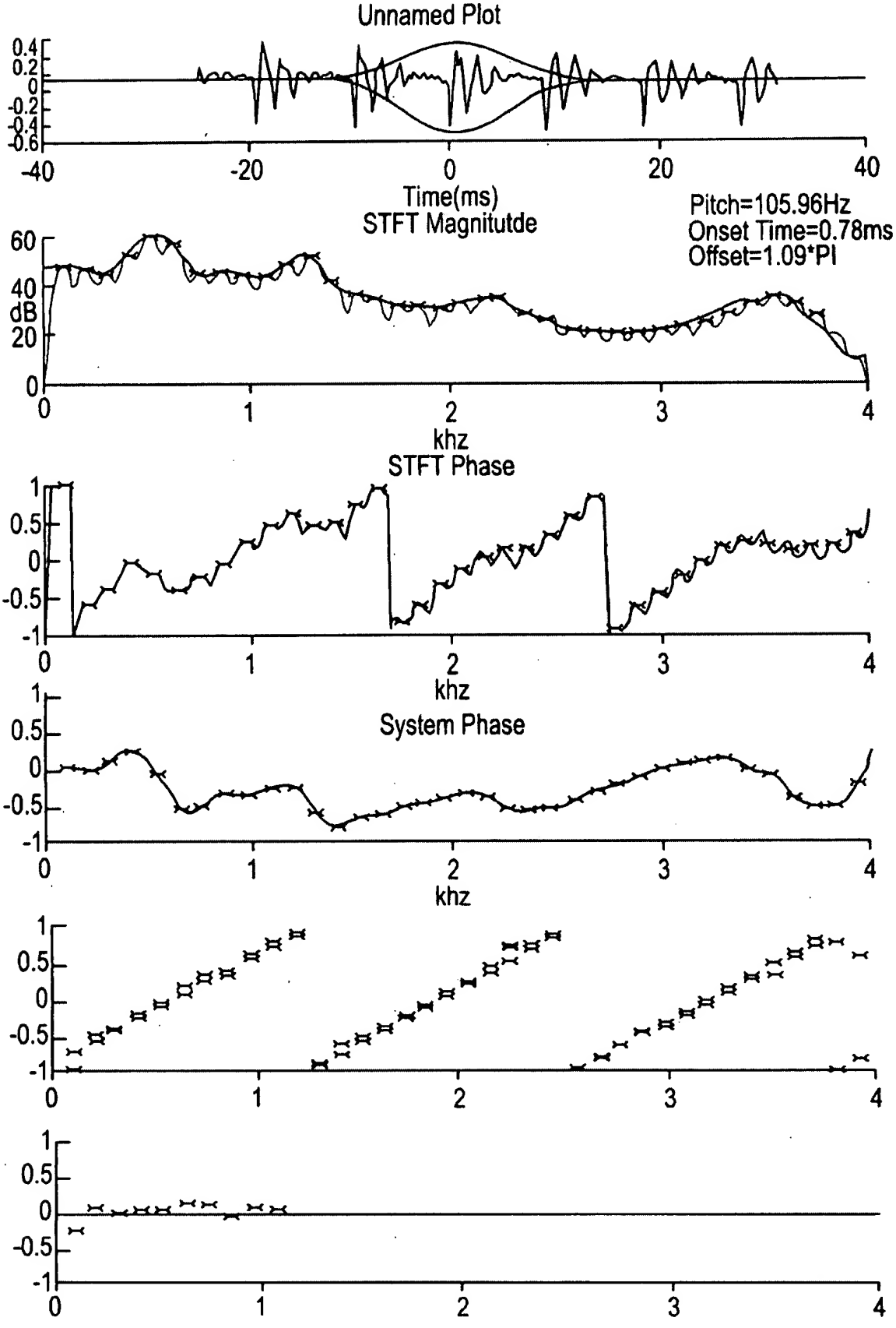

Unnamed Plot
0.4
0.2
0
-0.2
-0.4
-0.6
-40
-20
0
20
40
Time(ms)
STFT Magnitutde
Pitch=105.96Hz
Onset Time=0.78ms
Offset=1.09*PI
60
40
dB
20
0
0
1
2
3
4
khz
STFT Phase
1
0.5
0
-0.5
-1
0
1
2
3
4
khz
System Phase
1
0.5
0
-0.5
-1
0
1
2
3
4
khz
1
0.5
0
-0.5
-1
0
1
2
3
4
1
0.5
0
-0.5
-1
0
1
2
3
4

FIG. 20

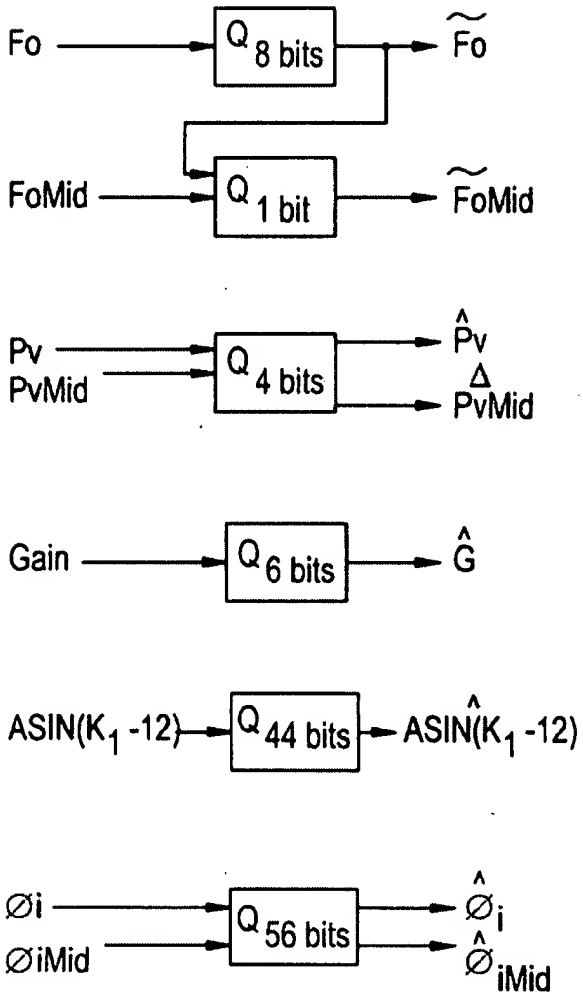

Fo
Q 8 bits
$\widetilde{Fo}$
FoMid
Q 1 bit
$\widetilde{FoMid}$
Pv
PvMid
Q 4 bits
$\hat{Pv}$
Δ PvMid
Gain
Q 6 bits
$\hat{G}$
ASIN(K1 -12)
Q 44 bits
$\hat{ASIN}$(K1 -12)
Øi
ØiMid
Q 56 bits
$\hat{Ø}_i$
$\hat{Ø}_{iMid}$

FIG. 21

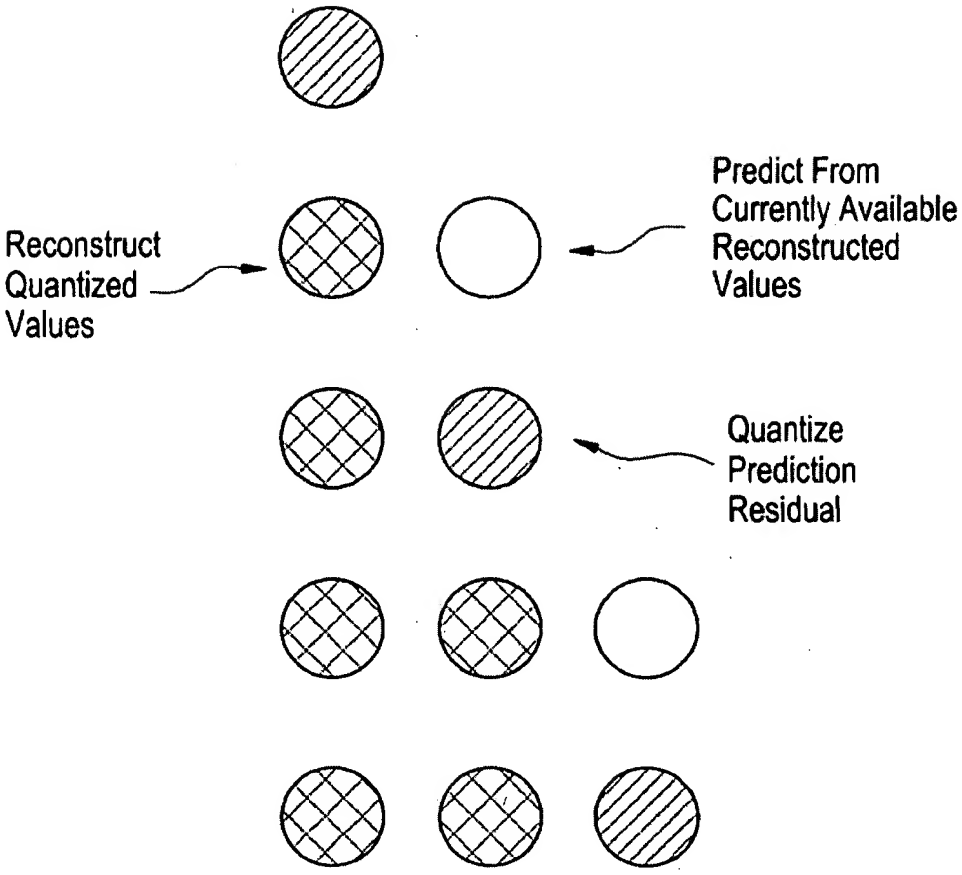
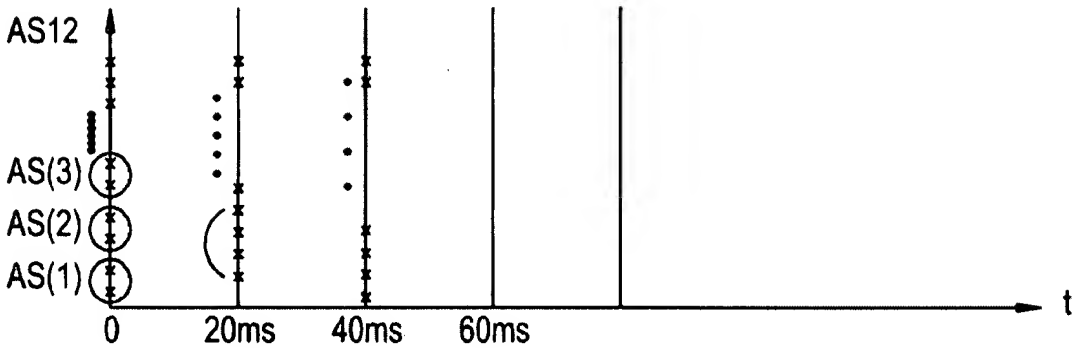

Predict From
Currently Available
Reconstructed
Values
Reconstruct
Quantized
Values
Quantize
Prediction
Residual

FIG. 21A



Prediction of AS(i)

$$\hat{AS}(i) = \sum_{j=1}^{i-1} \hat{a}_{ij} = \tilde{AS}(j)$$

$$\begin{bmatrix} \hat{AS}(1) \\ \hat{AS}(2) \\ \vdots \\ \hat{AS}(12) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ a_{21} & 0 & & & \\ a_{31} & a_{32} & & & \\ \vdots & & & & \\ a_{12,1} & a_{12,2} & \cdots & a_{12,11} & \end{bmatrix} \begin{bmatrix} \tilde{AS}(1) \\ \tilde{AS}(2) \\ \vdots \\ \tilde{AS}(11) \end{bmatrix}$$

12x1 — 12x11 — 1x1

Fixed Store

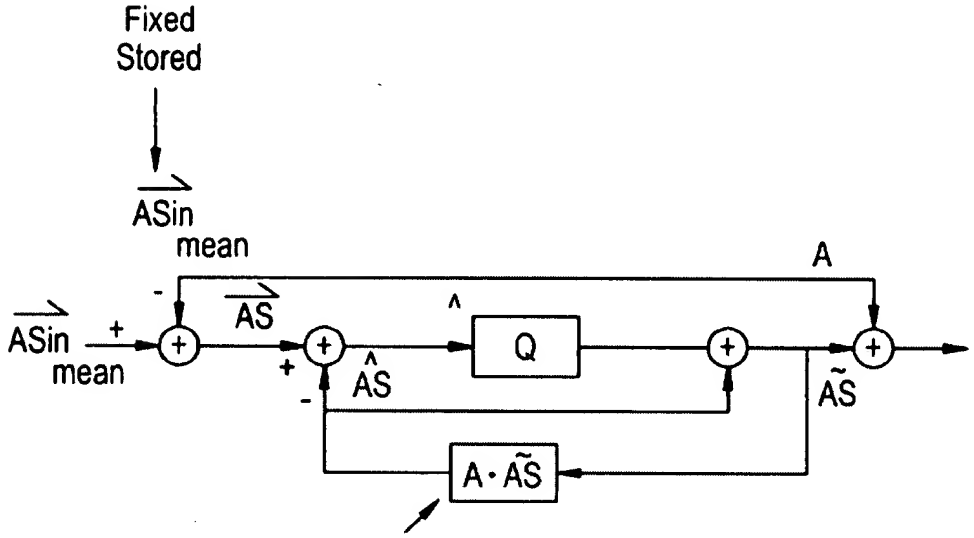
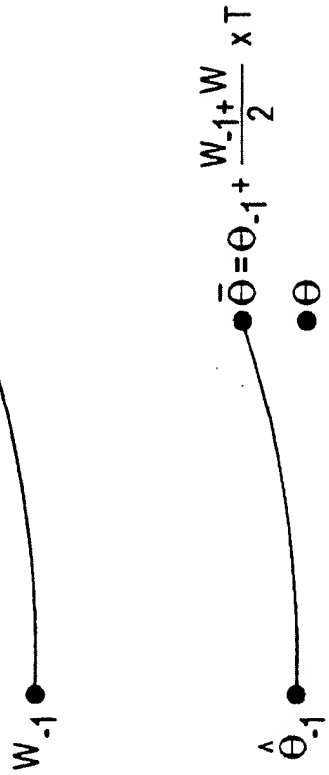


FIG. 22A



Phase Redidual = $\Theta - \bar{\Theta}$

W_{-1} = Frequency at Previous Frame

W = Frequency at Current Frame

Θ_{-1} = Quantized Phase at Previous Frame

$\bar{\Theta}$ = Predicted Phase at Current Frame

Θ = Measured Phase at Current Frame

FIG. 22B

Scatter Plot of 20ms Phase and 10ms Phase Prediction Error

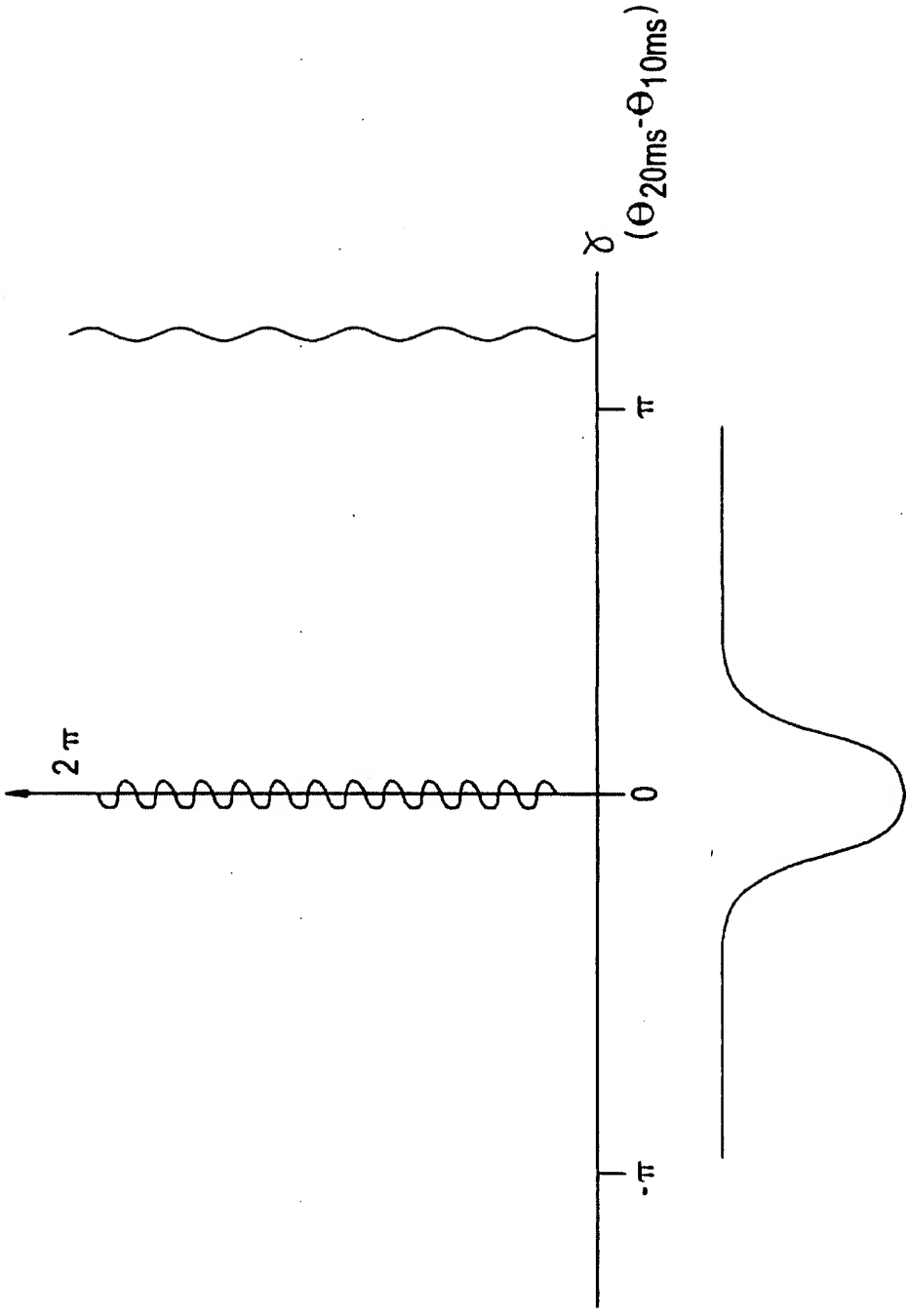


FIG. 23A

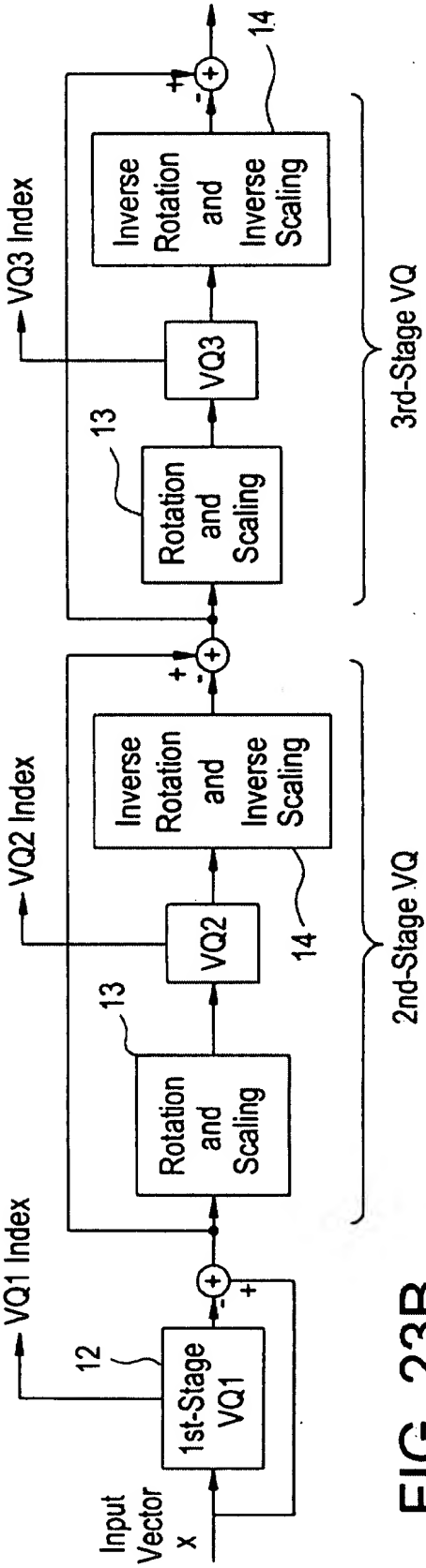
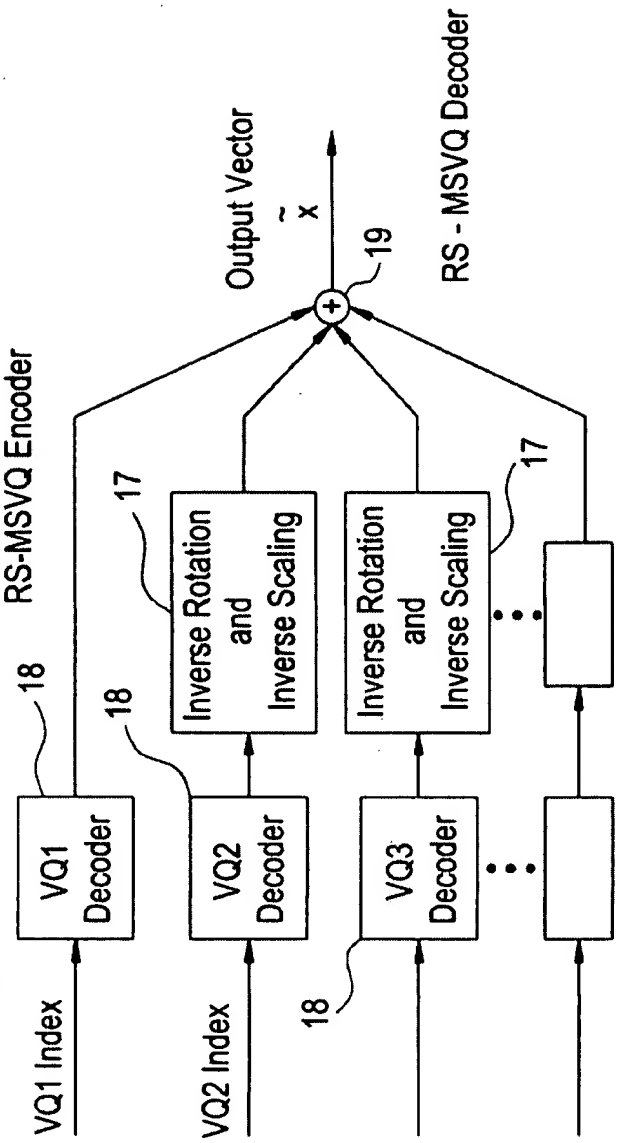
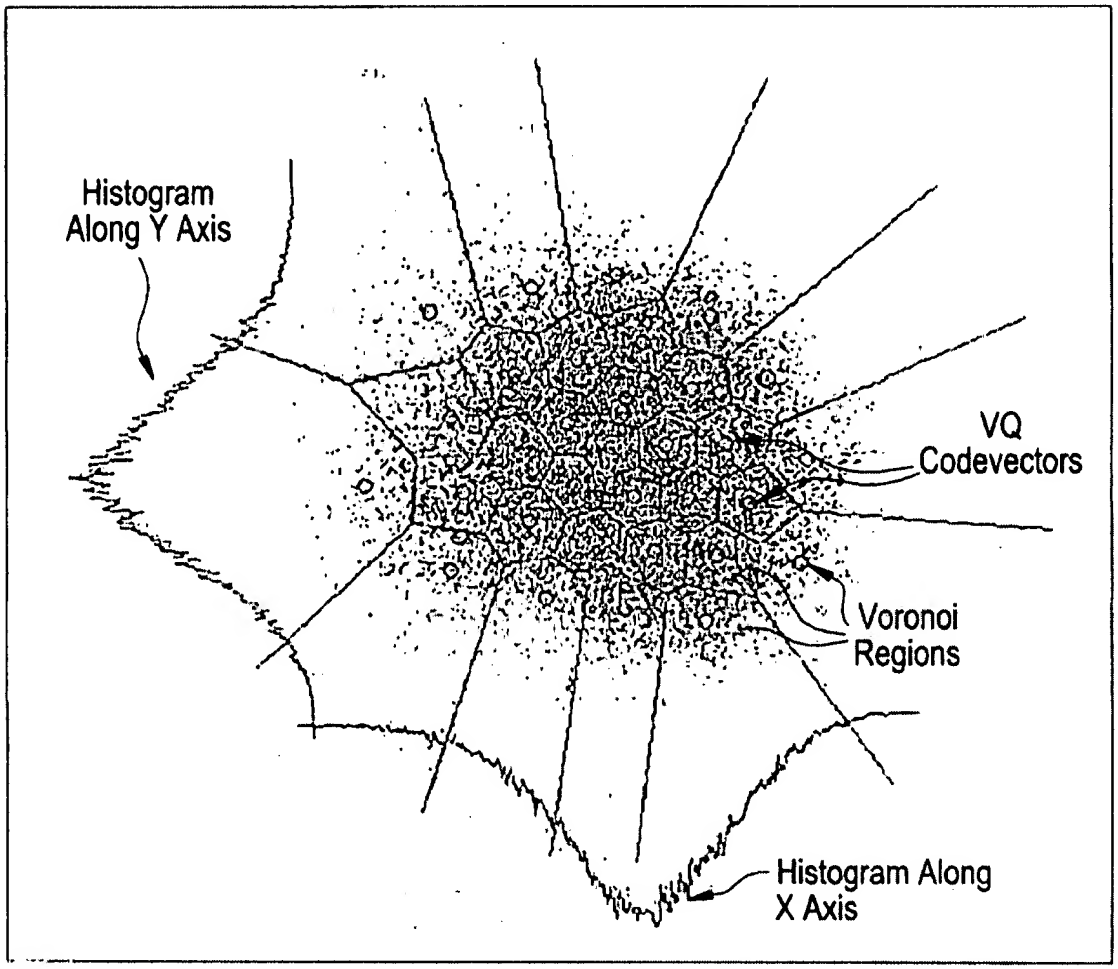


FIG. 23B



REPLACEMENT SHEET

FIG. 24A



Scatter plot (gray dots) of 4th pair of ASIN(dc) intra-frame prediction error, the histogram along each direction, and the corresponding 1st-stage 5-bit VQ codebook and Voronoi regions.

FIG. 24B

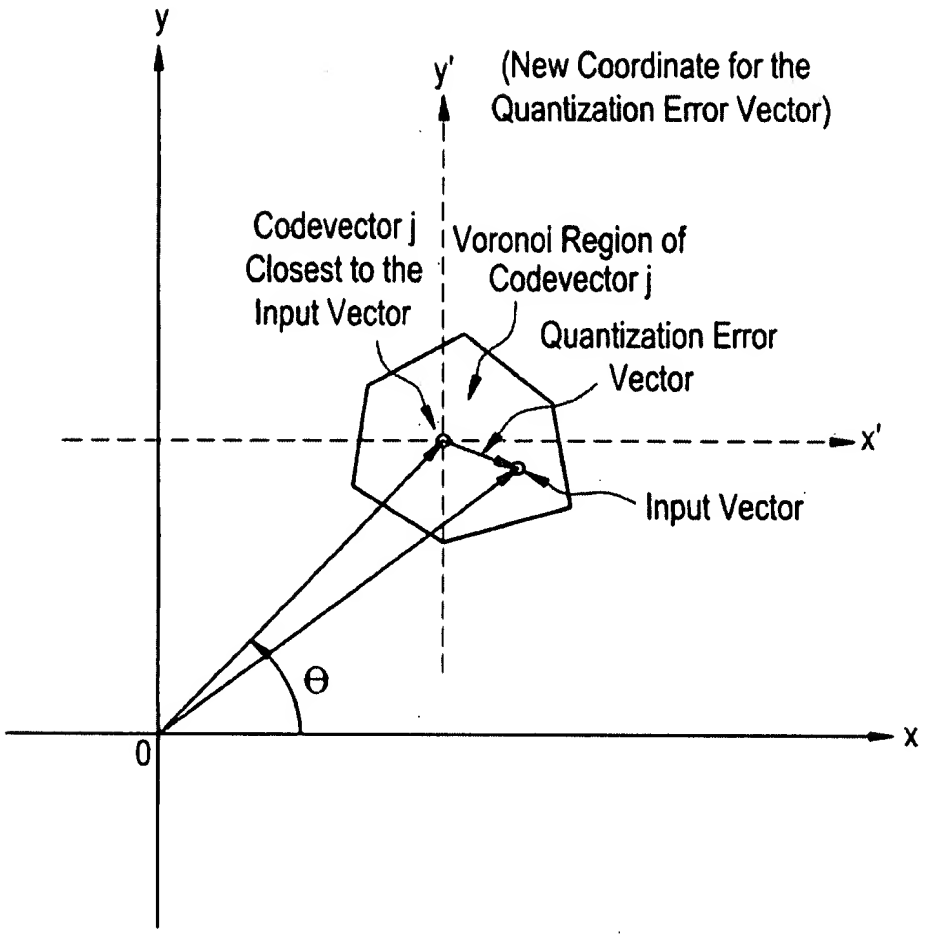
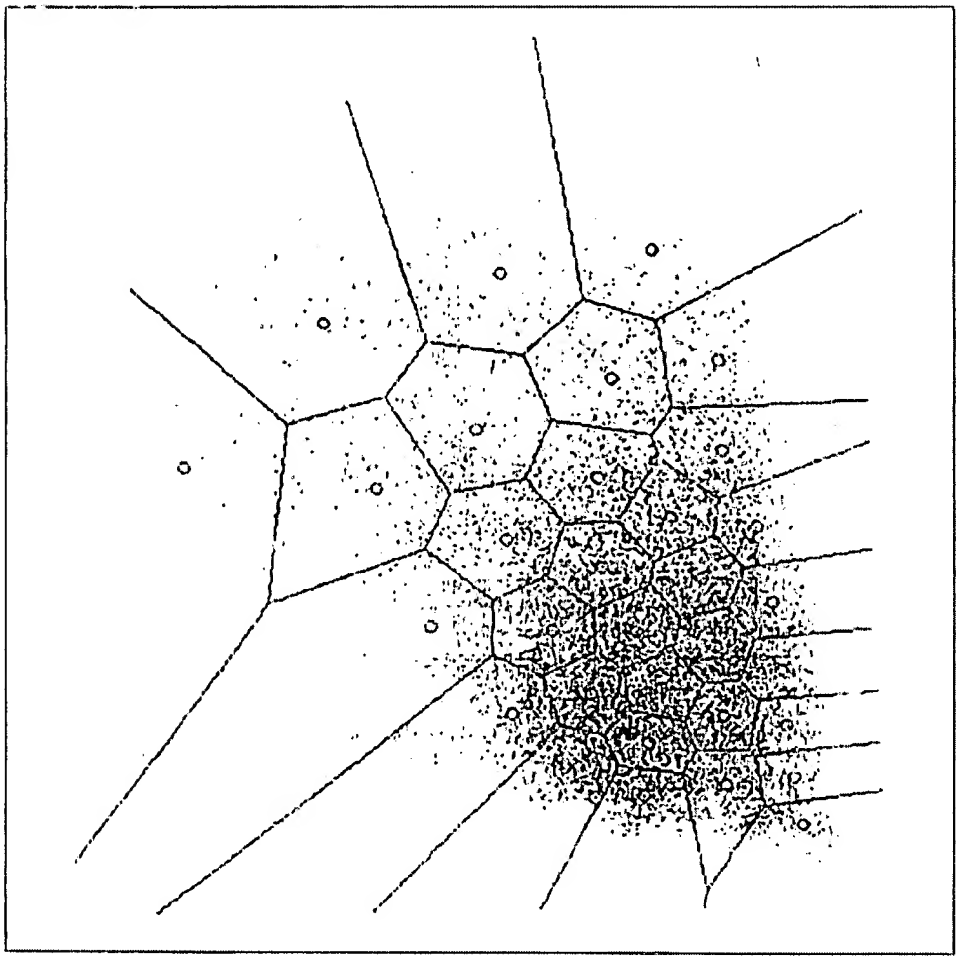
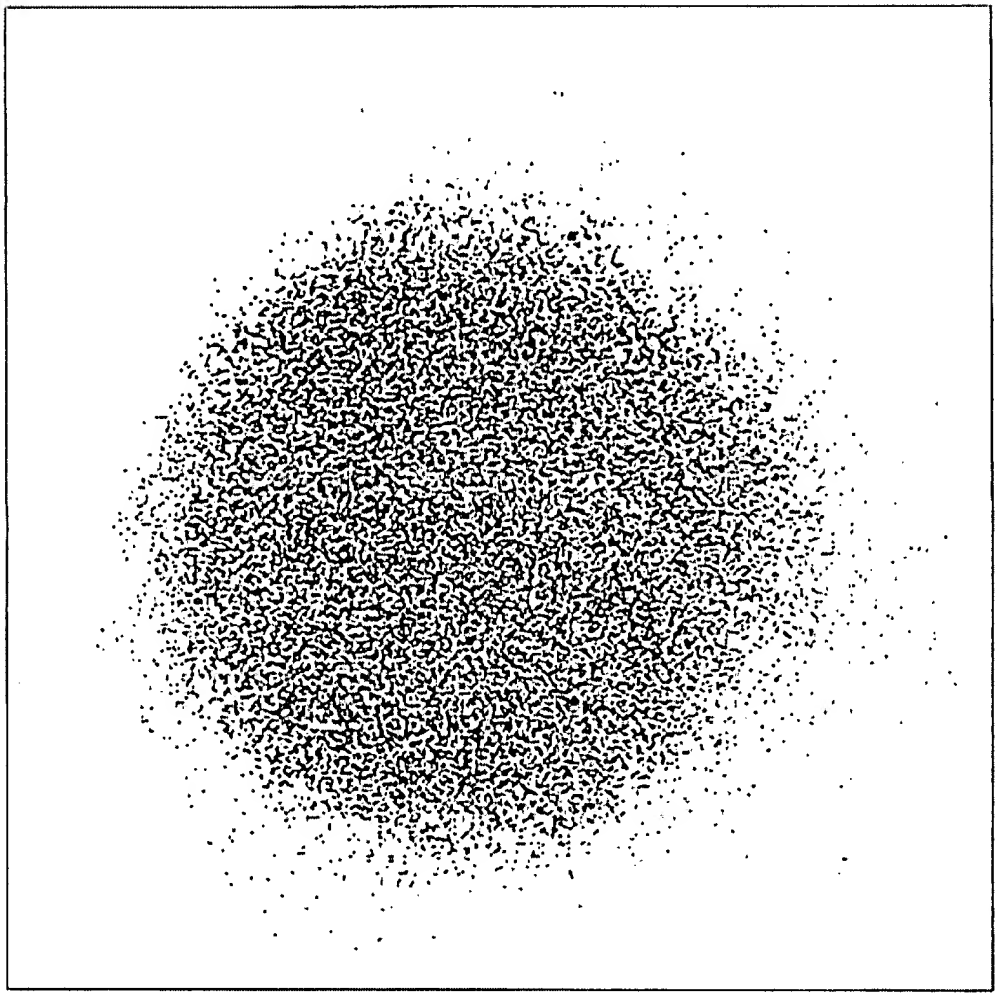

y
y'
(New Coordinate for the
Quantization Error Vector)
Codevector j
Closest to the
Input Vector
Voronoi Region of
Codevector j
Quantization Error
Vector
x'
Input Vector
θ
0
x

FIG. 24C



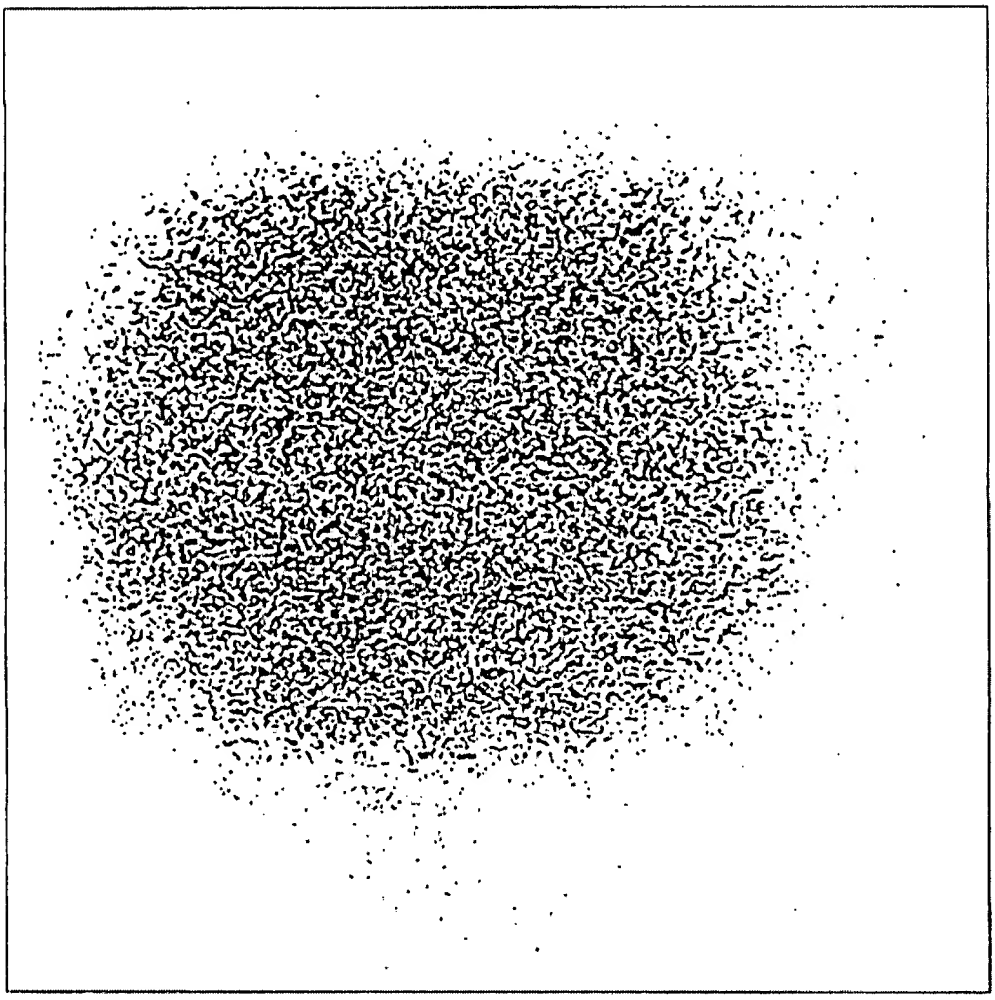
Scatter plot of 1st pair of ASIN(k) (gray dots) and
1st-stage VQ codebook (small circles) and the
corresponding voronoi cells

FIG. 25



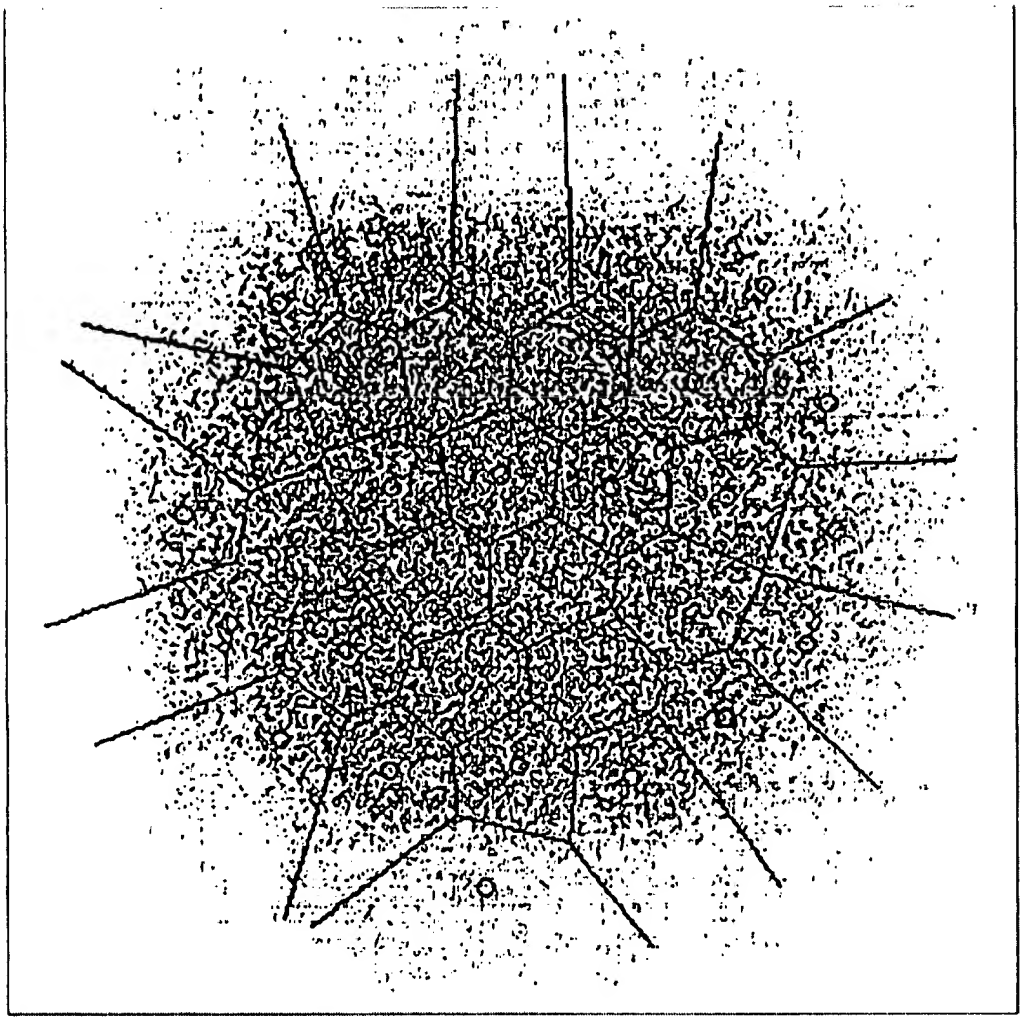
Without hand-tuned rotation angles
inner cells 1st-stage VQ of 1st pair of ASIN(k)

FIG. 26



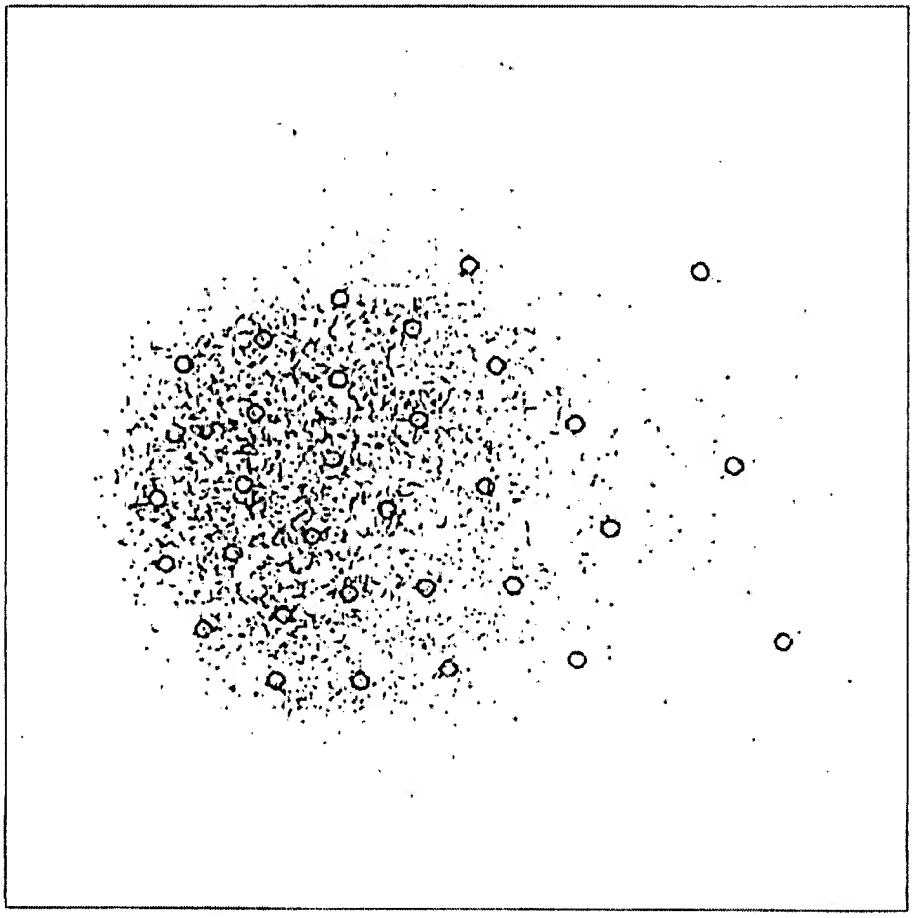
With hand-tuned rotation angles
inner cells of 1st-stage VQ of 1st pair ASIN(k)

FIG. 27



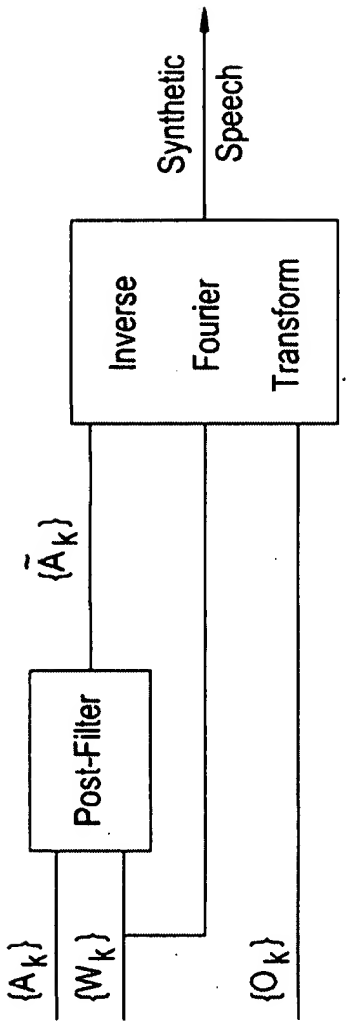
Inner-cell 1st-stage VQ error vector distribution (gray dots) (hand tuning) and corresponding 2nd stage VQ codebook (small circles) for 1st pair of ASIN(k)

FIG. 28



Outer-cell 1st-stage VQ error vector distribution
and corresponding 2nd-stage VQ codebook (small circle)
for 1st pair of ASIN(k)

FIG. 29



$\{A_k\}$ = k^{th} Sine-Wave Amplitude

W_k = k^{th} Sine-Wave Frequency

O_k = k^{th} Sine-Wave Phase

$\tilde{A}_k$ = k^{th} Post-Filtered Sine-Wave Amplitude

FIG. 30

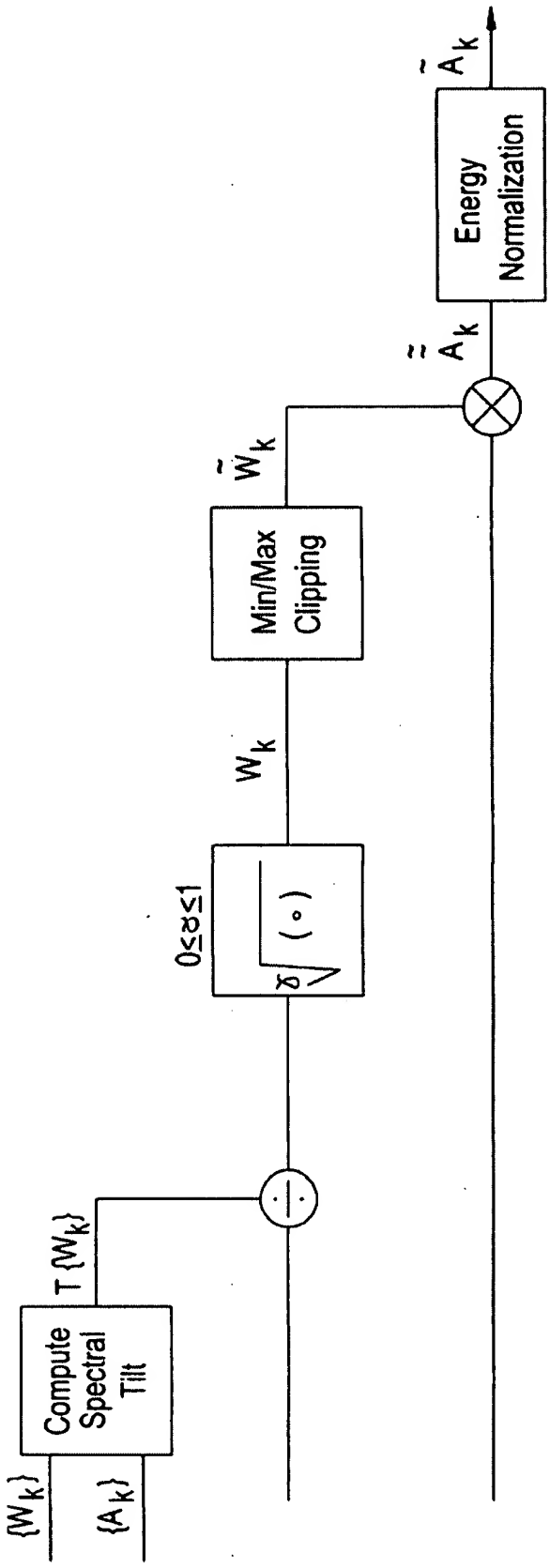
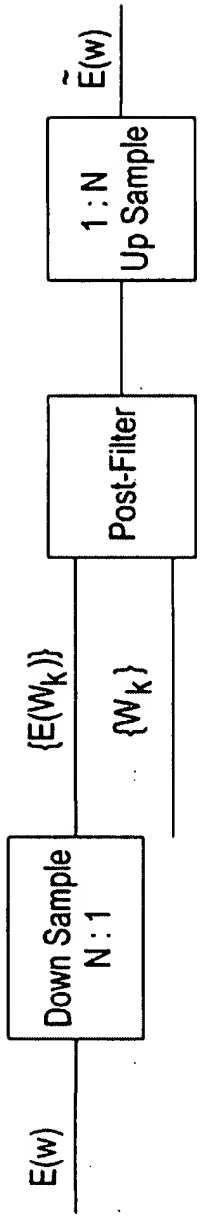

$\tilde{A}_k$
Energy Normalization
$\tilde{\tilde{A}}_k$
$\tilde{W}_k$
Min/Max Clipping
W_k
$0 \le \gamma \le 1$
$\gamma(\cdot)$
$T\{W_k\}$
Compute Spectral Tilt
$\{W_k\}$
$\{A_k\}$

FIG. 31



E(w) = Amplitude Envelope

W_k = Down Sampled Frequencies

$\tilde{E}(w)$ = Post-Filtered Amplitude Envelope

FIG. 32

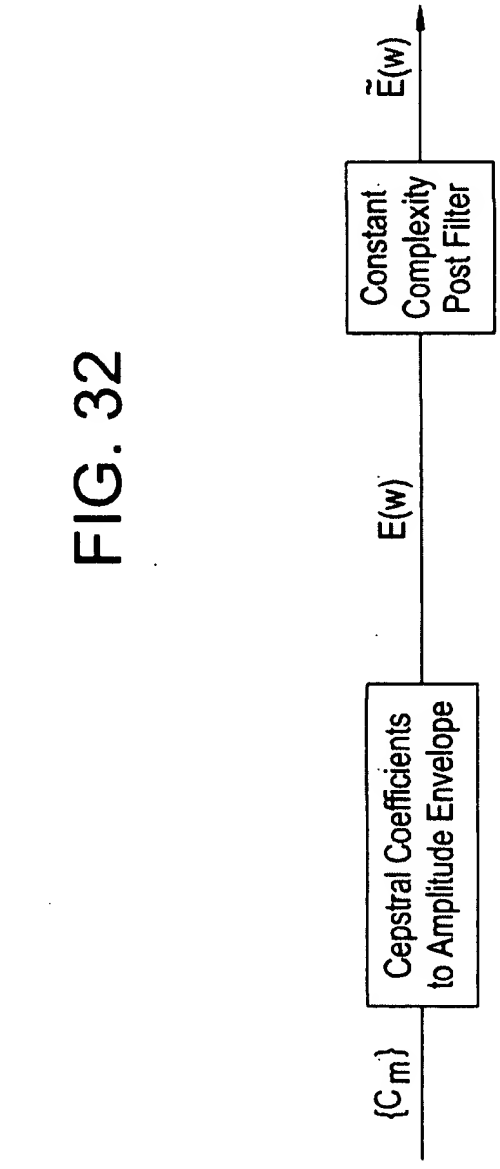

{Cm}
Cepstral Coefficients to Amplitude Envelope
E(w)
Constant Complexity Post Filter
Ẽ(w)
Cm = mth Cepstral Coefficient
E(w) = Amplitude Envelope
Ẽ(w) = Post-Filtered Amplitude Envelope

FIG. 33

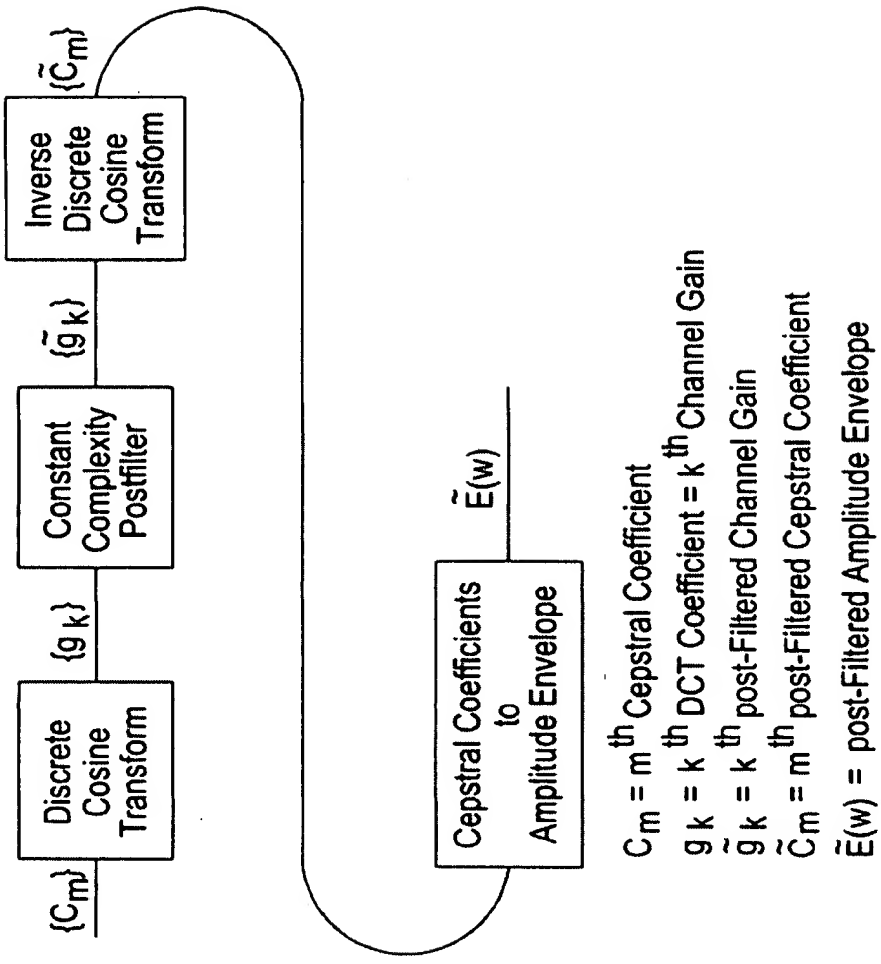

{Cm}
Discrete Cosine Transform
{gk}
Constant Complexity Postfilter
{g̃k}
Inverse Discrete Cosine Transform
{C̃m}
Cepstral Coefficients to Amplitude Envelope
Ẽ(w)
Cm = mth Cepstral Coefficient
gk = kth DCT Coefficient = kth Channel Gain
g̃k = kth post-Filtered Channel Gain
C̃m = mth post-Filtered Cepstral Coefficient
Ẽ(w) = post-Filtered Amplitude Envelope

FIG. 34

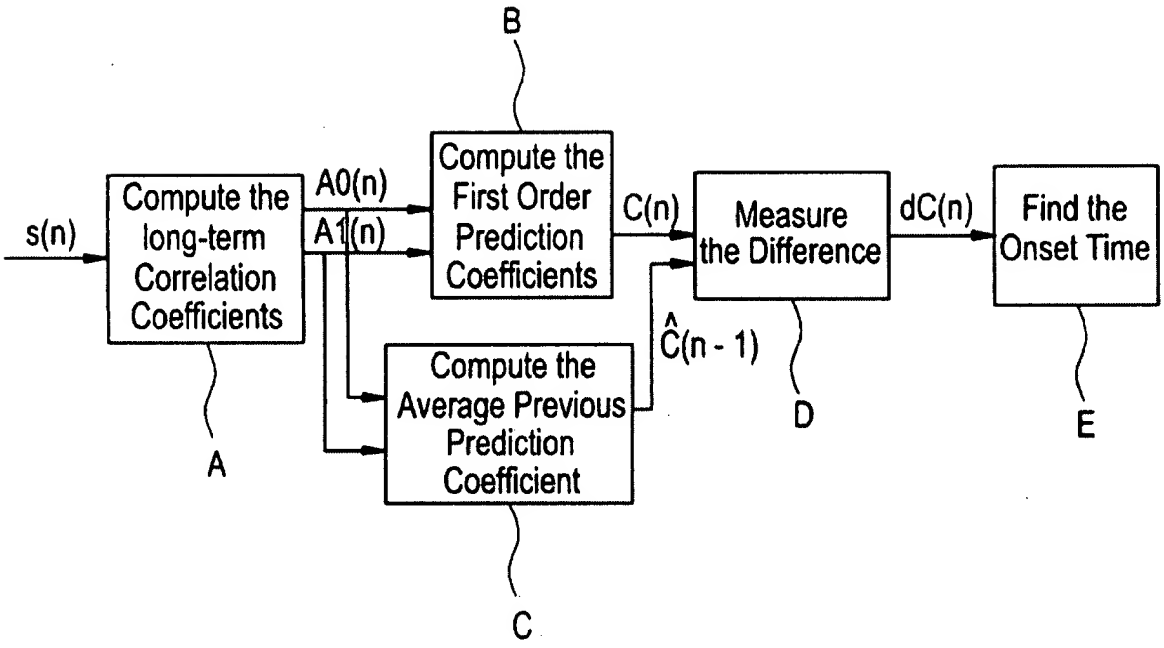
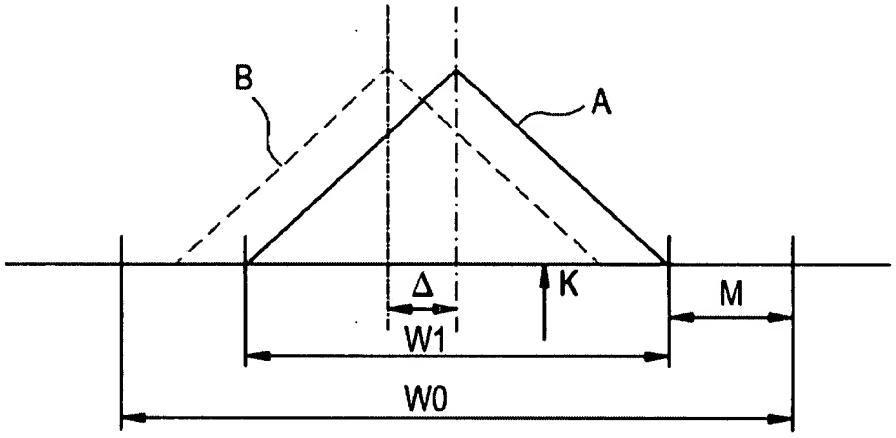

B
Compute the First Order Prediction Coefficients
Compute the long-term Correlation Coefficients
s(n)
A0(n)
A1(n)
C(n)
Measure the Difference
dC(n)
Find the Onset Time
Ĉ(n - 1)
Compute the Average Previous Prediction Coefficient
A
C
D
E

FIG. 35


B
A
Δ
K
M
W1
W0